



FEDERAL INFORMATION  
PROCESSING STANDARDS PUBLICATION  
1974 DECEMBER 1

U.S. DEPARTMENT OF COMMERCE / National Bureau of Standards



# OPTICAL CHARACTER RECOGNITION CHARACTER SETS

CATEGORY:      HARDWARE STANDARD  
SUBCATEGORY:    CHARACTER RECOGNITION

JK  
468  
A8A3  
no. 32  
1974

23 DEC 1974

NOT REC.

JK 468

A8A3

10-32

1974

C.2

## Foreword

The Federal Information Processing Standards Publication Series of the National Bureau of Standards is the official publication relating to standards adopted and promulgated under the provisions of Public Law 89-306, and Part 6 of Title 15 Code of Federal Regulations. The entire series constitutes the FEDERAL INFORMATION PROCESSING STANDARDS REGISTER.

The series is used to announce Federal Information Processing Standards, and to provide standards information of general interest and an index of relevant standards publications and specifications. Publications that announce adoption of standards provide the necessary policy, administrative, and guidance information for effective standards implementation and use. The technical specifications of the standard are usually attached to the publication, otherwise a reference source is cited.

Comments covering Federal Information Processing Standards and Publications are welcomed, and should be addressed to the Associate Director for ADP Standards, Institute for Computer Sciences and Technology, National Bureau of Standards, Washington, D.C. 20234. Such comments will be either considered by NBS or forwarded to the responsible activity as appropriate.

Richard W. Roberts, *Director*

## Abstract

This standard provides the description, scope, and identification for standard sets of graphic shapes to be used in the application of Optical Character Recognition (OCR) systems. Two font styles, known as Style A and B, are described. Style A comprises a font of 92 characters which is designed to provide a maximum of machine efficiency in reading under a wide variety of applications. Style B comprises a font of 96 characters, which stresses esthetic appearance, but which may be applied under a substantial range of applications. Three sizes of characters designated as Size I, III, and IV are presented. The basic requirements related to character positioning are also specified. Individual character drawings for both styles of character sets are included.

**Key Words:** Alternate character; centerline drawings; character positioning; character sets; character shape; character sizes; font; lower case character; Optical Character Recognition; upper case character.

U.S. Nat. Bur. Stand. (U.S.), Fed. Info. Process. Stand. Publ. (FIPS PUB) 32, 77 pages (1974)

CODEN: FIPPAT

For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402. (Order by SD Catalog No. C13.52:32.) Price \$1.40. Subscription service also available for all new FIPS publications and supplements for an indefinite period.



## Federal Information Processing Standards Publication

Date 1974 December 1



### ANNOUNCING THE STANDARD FOR OPTICAL CHARACTER RECOGNITION CHARACTER SETS

Federal Information Processing Standards Publications are issued by the National Bureau of Standards pursuant to the Federal Property and Administrative Services Act of 1949 as amended, Public Law 89-306 (79 Stat. 1127), and as implemented by Executive Order 11717 (38 FR 12315, dated May 11, 1973), and Part 6 of Title 15 CFR (Code of Federal Regulations).

**Name of Standard.** Optical Character Recognition Character Sets.

**Category of Standard.** Hardware Standard, Character Recognition.

**Explanation.** This standard provides the description, scope, and identification for standard sets of graphic shapes to be used in the application of Optical Character Recognition (OCR) systems.

**Approving Authority.** Secretary of Commerce.

**Maintenance Agency.** Department of Commerce, National Bureau of Standards (Institute for Computer Sciences and Technology).

#### Cross Index.

a. ANSI X3.17-1974 (Revised), American National Standard Character Set for Optical Character Recognition.

b. ECMA 11 (Revised), European Equipment Manufacturers Association Standard Character Set for Optical Character Recognition.

c. ISO 1073, International Standard for Alphanumeric Character Sets for Character Recognition.

**Applicability.** This standard is applicable to Optical Character Recognition systems utilizing any part or all of the character sets contained herein. This standard provides for two different character sets (OCR-A and OCR-B). The selection of which of these sets to utilize is a decision to be made based upon the operational requirements of specific applications.

**Implementation Schedule.** All applicable equipment ordered on or after the date of this FIPS PUB must be in conformance with this standard unless a waiver has been obtained in accordance with the procedure described below. Exceptions to this standard are made in the following cases:

- a. For equipment installed or on order prior to the date of this FIPS PUB.
- b. Where procurement actions are into the solicitation phase (i.e., Request for Proposals or Invitation for Bids have been issued) on the date of this FIPS PUB.

**Waiver Procedure.** Heads of agencies may waive the provisions of the implementation schedule. Proposed waivers relating to procurement of nonconforming equipment or the use of nonconforming character sets will be coordinated in advance with the National Bureau of Standards. Letters should be addressed to the Associate Director for ADP Standards, Institute for Computer Sciences and Technology, National Bureau of Standards, Washington, D.C. 20234. They should describe the nature of the waiver and set forth the reasons therefor.

Sixty days should be allowed for review and response by the National Bureau of Standards. The waiver is not to be effective until a reply is received from the National Bureau of Standards; however, the final decision for the granting of a waiver is a responsibility of the agency head.

**Specifications.** Federal Information Processing Standard 32, Optical Character Recognition Character Sets (affixed).

**Qualifications.** As contained in the specifications.

**Where to Obtain Copies of the Standard.** Copies of this publication are for sale from the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402, SD Catalog Number C13.52:32). There is a 25 percent discount on quantities of 100 or more. When ordering, specify document number, title, and SD Catalog Number or Accession Number.





## Federal Information Processing Standard 32

Date 1974 December 1



### SPECIFICATIONS FOR OPTICAL CHARACTER RECOGNITION CHARACTER SETS

1. **Name of Standard.** Optical Character Recognition Character Sets.
2. **Category of Standard.** Hardware Standard, Character Recognition.
3. **Explanation.** This standard provides the description, scope, and identification for character sets of graphic shapes to be used in the application of Optical Character Recognition Systems.
4. **Specifications.** Optical Character Recognition (OCR) Character Sets are designated collections of graphic shapes ordered into full sets and subsets to be used in the application and operation of OCR systems between and among agencies.

4.1. Character sets of varying characteristics are provided to meet the several levels of requirements of user agencies. Initial character sets are designated as:

Part II—Style A  
Part III—Style B

Other character sets may be developed and added to this specification from time to time.

4.2. Specific characteristics of the above character sets are contained in separate parts which follow hereafter.

5. **Qualifications.** A family of related standards is required in order to describe the full set of performance characteristics necessary for a complete operational OCR system. This standard is a member of this family. Other standards will cover OCR Forms and OCR Print Quality. Additionally, guidelines will be prepared to facilitate the use of OCR and standard implementation.

6. **Special Information.** In general, the principal features of the character sets listed above are as follows:

**Style A**—A font of 92 characters which is designed to provide a maximum of machine efficiency in reading under a wide variety of applications. Subsets are provided for situations in which less than the full repertoire is indicated. The set maps into the ASCII Code Table (FIPS 1).

**Style B**—A font of 96 characters, which stresses esthetic appearance but which may be applied under a substantial range of applications. Care must be exercised to provide for the high level of print quality necessary for efficient machine performance. Subsets are provided for situations where less than the full repertoire is indicated. The set maps into the ASCII Code Table (FIPS 1).

## FIPS 32

In general, the character sets of Part II and Part III follow the American National Standard X3.17 Revised and ECMA 11 Revised. Correspondence with these standard documents will be maintained to the extent possible, in the ongoing maintenance of this FIPS standard.

The inch and metric dimensions used in this FIPS PUB are not precisely equivalent. To achieve consistency, the two sets of dimensions should not be intermixed.

## PART I—GENERAL

### 1. Introduction

1.1. **Scope.** Parts II and III of this FIPS define the shapes and sizes of characters contained in Optical Character Recognition (OCR) Character Sets—Style A and Style B—and establish specifications and recommendations for:

- a. the optical and dimensional properties of the shape patterns forming OCR characters;
- b. the basic requirements related to the position of OCR characters on the paper substrate.

1.2. **Purpose.** The purpose of Parts II and III of this FIPS is to establish standard character sets to be used in OCR systems and to aid in the implementation and use of such systems. The character repertoires include a lower case alphabet and a CHARACTER ERASE symbol, primarily intended to be used on typewriters, and designed to be recognized intermixed with the remainder of the characters. A GROUP ERASE symbol is provided for use either with typewriters or in a later proof reading step.

The character sets of Part II and Part III are taken from U.S. and European National Area Standards. Four sizes of characters were originally postulated, Sizes I, II, III, and IV. With the passage of time the series contemplated for Size II is no longer used and the designation is no longer meaningful. In order to maintain correspondence with existing national standards X3.17 and ECMA-11 the designations of Sizes I, III, and IV will be used in this FIPS PUB.

1.3. **Use of the Standard.** An OCR system must detect printed characters by means of differences in the reflected light. These differences are detected by a system of one or more electronic photodetectors associated with optical and mechanical apparatus.

These optical readers lack the versatility of the human visual system and usually have some distinctively unique characteristics. OCR readers are typically responsive to different wavelength bands of light (or colors) than the human eye. Some readers may be responsive to wavelengths outside of the visual range. In general, an OCR reader lacks the discrimination of the eye with respect to difference in color contrast between a printed image and its background. The humanly visual appearance may be misleading unless the parameters of the reader are understood and taken into account. The various provisions which follow are related to these machine based parameters and characteristics.

## PART II—STYLE A

### 2. Standard Characters

**2.1. Character Sizes.** Standard character shapes are specified in three different sizes (I, III, & IV) (Figures II-11 through II-67), except for the lower case letters (Figures II-68 through II-96) which are specified only in the smallest size (Size I). Table I below specifies the basic centerline dimensions (W and H) of these three sizes. It also indicates the nominal strokewidth T and the minimum length L of Long Vertical Mark (LVM).

TABLE II-1—Nominal Character Sizes—Style A

Size	Nominal centerline width (W)		Nominal centerline height (H)		Nominal stroke Width (T)		Minimum length of long vertical mark (L)	
	<i>Inch</i>	<i>(mm)</i>	<i>Inch</i>	<i>(mm)</i>	<i>Inch</i>	<i>(mm)</i>	<i>Inch</i>	<i>(mm)</i>
I.....	0.055	(1.40)	0.094	(2.40)	0.014	(0.35)	0.146	(3.73)
III.....	0.060	(1.52)	0.126	(3.20)	0.015	(0.38)	0.196	(4.98)
IV.....	0.080	(2.04)	0.150	(3.80)	0.020	(0.51)	0.233	(5.91)

**Long Vertical Mark.** There is no specified maximum length for the Long Vertical Mark (Figure II-24). It may, for example, be a ruled line from top to bottom of the form. The LVM shall in every case, except for those lower case characters that have descenders, extend beyond the highest and lowest portion of any character in a printed line.

**Lower Case Characters.** The lower case characters i, j, m, p, and w were designed to exceed the nominal values given in Table I. (See Figures II-76, 77, 80, 83, and 90).

**Inch/Metric Equivalents.** The inch and metric dimensions in this standard are not precisely equivalent. For purposes of consistency, type designers should adopt the use of either system but not intermix them.

**Erase Characters.** CHARACTER ERASE and GROUP ERASE were designed to exceed the nominal values because of their unique characteristics (See Figures II-95 and II-96).

**2.1.1. Character Set Application.** The three sizes of characters for Style A are nominally applied as follows:

**Size I** was developed for devices such as high-speed line printers and typewriters.

**Size III** was developed to meet the requirements of printers such as cash registers and accounting machines.

**Size IV** was developed to meet the requirements of printing from embossed plastic cards and metal plates.

Nothing in the preceding statements is intended in any way to limit the applications of any of the sizes to any particular printing device, but simply to caution the user that due consideration should be given to the selection of font size for a given application.

**2.2. Character Set Repertoire.** The printing graphics and character SPACE as defined in this standard constitute the total repertoire for Optical Character Recognition for data input purposes. In various systems applications it may be desirable to use special characters herein defined, such as FORK, HOOK, and CHAIR, for the purposes of error correction, device control or similar nondata functions.

The OCR character set described in Figures II-11 through II-96 comprises those graphics which are most commonly used by the data processing industry. The overall number of graphics, therefore, may be greater than that required for a particular application. Since large character sets may adversely affect printer throughput and machine-recognition performance, it is recommended that a full consideration of font requirements be made to attain optimum OCR system operation.

2.2.1. There are no restrictions as to the information content of any of the standard OCR characters except for CHARACTER ERASE, GROUP ERASE, and SPACE. The meaning of any characters used in any particular application must be established by the user. Users are cautioned to ensure that a common understanding of the character sets employed in applications involving interchange of documents has been established.

2.2.2. **Lower Case Alphabet.** The 26 lower case alphabetic letters are intended primarily for use on manually operated serial entry devices (i.e., typewriters) to accommodate needs for a compatible and scannable lower case alphabet. Users must recognize that some OCR reader systems may not have the capability of reading the lower case alphabet.

2.3. **Subsets.** Subsets are NOT defined herein. They are the subject of a separate FIPS PUB. The user is cautioned that the complete repertoire may not be necessary, and an expanded set may adversely affect system performance. It is recommended that an appropriate minimum set be selected for each application.

2.3.1. **Character Shapes to be Developed.** Graphic shapes for the following characters are under development and will be available for implementation at a later date—

GREATER THAN	OPENING PARENTHESIS
LESS THAN	CLOSING PARENTHESIS
REVERSE SLANT	OPENING BRACKET
NUMBER SIGN	CLOSING BRACKET
EXCLAMATION POINT	COMMERCIAL AT

When these graphic shapes are available the present OPENING and CLOSING PARENTHESIS will be redesignated as OPENING and CLOSING BRACE.

2.3.2. **Alternate Character Shapes.** Alternate graphic shapes have been developed for PERIOD, COMMA, and QUESTION MARK for use in high speed printing applications where "print through" or punch through is a problem. The alternate graphic shapes avoid the use of small printing areas which may permit physical penetration of the form substrate. It is recommended that future OCR system designs accommodate both original (typewriter) and alternate (high speed printer) graphic shapes.

New shapes were developed for the HYPHEN and APOSTROPHE to overcome both recognition separability and punch through problems with the original shapes. The original shapes are now designated as alternates and should be recognized as causing a potential problem.

The alternate graphic shapes are illustrated in Figures II-97 through II-101.

2.3.3. **Optional Characters.** CHARACTER ERASE, GROUP ERASE, and Long Vertical Mark are optional characters which MAY be used with any subset or portion thereof and which MUST be read by any reader.



FORK, HOOK, and CHAIR are optional data characters which MAY be used with any subset or portion thereof and which MUST be read for any reader other than those which read the FIPS Basic Numeric Subset.

**2.4. Relationship to ASCII Code Table.** It should be noted that the ASCII Code Table (see FIPS PUB 1 and 15) includes the characters CIRCUMFLEX, UNDERLINE, GRAVE ACCENT, and OVERLINE (TILDE) which are not required in OCR applications, even with text that contain these accents. Surveys of interested user communities show no requirement for machine reading of these graphic shapes. Accordingly, no provisions have been made by the reading machine industry to supply graphic shapes for these ASCII Code Table entries.

The correspondence of the Optional Characters is handled as follows:

SPACE is a normally nonprinting graphic shape and corresponds exactly with the SPACE of Code Table Position 2/0.

CHARACTER ERASE and GROUP ERASE are format effectors in that the action of the reading machine is to ignore the character(s) with CHARACTER ERASE superimposed upon them and to eliminate the line space otherwise occupied by them. GROUP ERASE brings forth a similar action by the reading machine except that a group of characters is ignored (the reader may be programmed to search elsewhere on the form for the corresponding corrected entry). Since no characters directly relating to CHARACTER ERASE or GROUP ERASE are normally transmitted from the reader or its peripherals, no entry in the ASCII Code Table is appropriate or required. If the user must produce output coding for the CHARACTER ERASE or GROUP ERASE they shall be transmitted as DELETE in Code Table Position 7/15.

Long Vertical Mark is a graphic shape most generally associated with the function of field mark. It is usually used to denote the limits of fields or data elements on OCR forms, particularly in applications in which data is entered with keyboard driven devices. LVM can be associated for data transmission purposes, if this is desired, with the ASCII Character Vertical Line in Code Table Position 7/12.

HOOK, CHAIR, and FORK are special characters or abstract symbols usually associated with machine instructions. As a general rule Long Vertical Mark, FORK, HOOK, and CHAIR should not appear in output data; although they can be used as control or information symbols. They can and have been used in OCR applications to contain data content, however, and hence could be used as a transmittable character. For this reason, HOOK, FORK, and CHAIR can be coded into the ASCII Code Table as replacements for Underline (Position 5/15), Grave Accent (Position 6/0), and Tilde (Position 7/14), respectively.

Because of possible conflicts with the alphabetic symbol Y, the use of the fork in alphabetic applications is not recommended.

**2.5. Character SPACE.** The character SPACE is a blank area in a print line having a width equal to the width of the character pitch. When a blank area is bounded by narrow characters, the characters shall be assigned a width of  $W + T$  (See Section 2.1) for purposes of determining the number of SPACE characters between the printed characters.

The accuracy with which the number of SPACE characters in a row can be determined depends upon the OCR reader used, the print location tolerances, and other factors. The width of multiple spaces and the response of a character reader to multiple spaces is, therefore, not covered in this FIPS PUB.

**2.6. CHARACTER ERASE and GROUP ERASE.** In many applications using manually operated serial entry devices (e.g., typewriters), it is advantageous to be able to correct errors as they are detected by the operators. Although error correction can be provided by many different methods, it would be desirable to provide a method of correction which requires neither additional space on the document nor post-editing of the data.

For this reason, and to facilitate information interchange, two symbols, CHARACTER ERASE and GROUP ERASE, are defined (see Figures II-95 and II-96).

CHARACTER ERASE is a normal full-size symbol which can be printed with a single stroke. It can be recognized distinctly with respect to other characters in this Part II solely by its unique (total) printing area. Other characters are distinguishable by using their printed (black) and unprinted (white) areas. This property of CHARACTER ERASE permits it to be recognized as it stands alone, or as printed over any other character in the standard. In addition, this permits CHARACTER ERASE to be recognized when printed over other nonstandard characters or printed images whose area is not larger than the area of CHARACTER ERASE.

GROUP ERASE is designed so that a long string of characters can be erased without striking a CHARACTER ERASE for each character to be deleted. It is defined as a continuous line between  $1/6$  H and  $5/6$  H above the nominal base line at least 0.300 in (7.62 mm) long, having a minimum thickness of 0.008 in (0.20 mm).

**2.7. Character Shapes and Dimensions.** The Figures II-11 through II-96 together with the dimensions given in Table II-3 define the characters in the standard set. The nominal printed image of each character is specified by its stroke centerlines and by its nominal strokewidth. It is recommended that the printed shapes also conform to the character outline shapes in Figures II-11 through II-101, that is, where sharp outline corners are shown, the image should likewise be as sharp as practical. However, it is recognized that some type making and printing processes will not be able to produce sharp corners, and it is not required that the printed image radii be less than 0.004 inch (0.1 mm). Note that upper case character dimensions are given in terms of W and H; lower case character dimensions are in thousandths of an inch, and are compatible with the Size I only.

### 3. Character Positioning

**3.1. Format Rules.** Character positioning specifications (format rules) are needed to ensure that each OCR character is seen by the reading device without interference from other OCR characters or from non-OCR matter. This section contains basic specifications relating to the position of characters on a form to accommodate general requirements of OCR devices. It does not contain all the rules which may be necessary for a particular application.

**3.2. Form Reference Edges.** Some specifications in this section relate to form reference edges. These can be horizontal and/or vertical edges. Because of the diverse nature of OCR forms, it may sometimes be convenient to specify one reference edge (e.g., for journal tapes); for others it may be necessary to specify two edges (e.g., for checks the bottom and right hand edges are usually specified). Character alignment is relative to these reference edges. (See Figure II-1.)

**3.3. Clear Area.** A clear area is defined as that region of a form reserved for the OCR characters and the clear space around these characters. OCR printing should be isolated from all other machine-detectable printing or patterns in order to allow the reading device to distinguish the OCR information more readily. The locations and dimensions of clear areas will

be determined by the nature of individual applications and the requirements specified in this section. This does not preclude the use of nonread inks for field titles suitable to the application within this area. (See Figure II-1.)

**3.4. Printing Area.** A printing area is a rectangle inside the clear area, in which only OCR characters are to be printed. The sides of this rectangle should be parallel or perpendicular to a form reference edge. The distances (a, b, c, d of Figure II-1) between the corresponding boundaries of the printing area and the clear area should not be less than 0.1 in (2.5 mm).

**3.5. Margin.** The distance between any boundary of the printing area and the nearest parallel form edge is called the margin. (See Figure II-1.)

A margin shall be at least 0.250 in (6.35 mm).

Where manually operated serial entry devices (e.g., typewriters) are used, the top and bottom margins shall be 1 in (25.4 mm).

There are special cases where the small size of the form may make large margins impractical and the boundary of the Printing Area may then have to lie close to the edge(s). Relaxation of the specification in this respect is permissible only when it has been established that all OCR devices in the system can handle such forms.

**3.6. Data Fields.** The concepts of lines and fields are often confused. For the purpose of this standard a data field is defined as specific portion of the printing area that is limited to sets of one or more characters that may be treated as a unit of information. These character sets may be located on one or more consecutive lines of printing. A line could comprise several fields. Dimensional specifications on fields do not appear in this standard.

**3.7. Line Boundary.** A line boundary (see Figure II-2) is defined as the smallest rectangle with sides parallel and perpendicular to a document reference edge, which contains all the boundaries of the component characters of the line.

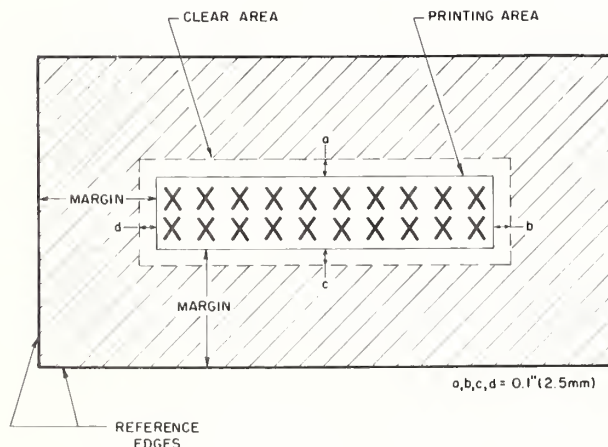


Figure II-1  
Margin Definition

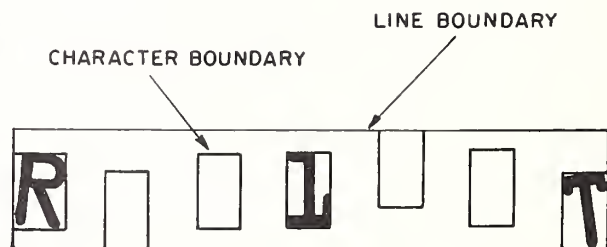


Figure II-2  
Character and Line Boundary

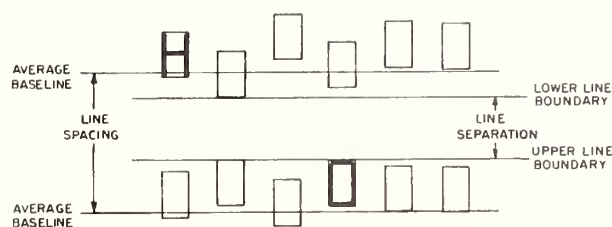


Figure II-3  
Line Spacing and Definition

**3.8. Line Spacing.** Line spacing (see Figure II-3) is the vertical distance between the average baseline position of all OCR characters printed on one line and that of all OCR characters printed on the next line. Nominal line spacing must be selected in such a way as to comply with the line separation tolerance. (The parameters which influence line separation are: line pitch, line skew, vertical misalignment, character height and strokewidth.)

Size	I	III	IV
Minimum line spacing.....	0.157 in (4.00 mm)	0.188 in (4.78 mm)	0.210 in (5.33 mm)
Nominal lines/inch.....	6	5	4

If character sizes are intermixed, the limitation applying to the largest size applies.

When lower case Size I characters are being used, there shall be no more than five lines/inch.

**3.9. Line Separation.** Line separation is the vertical distance between the upper line boundary (see Section 3.7) for a line of print, and the lower line boundary for the line immediately above (see Figure II-3).

Minimum line separation shall not be less than the following values:

Size	I	III	IV
Minimum line.....	0.025 in	0.060 in	0.080 in
Separation.....	(0.64 mm)	(1.52 mm)	(2.03 mm)

The line separation should be maintained as large as possible by means of a reduction in vertical misalignment of the characters and by close conformity to the nominal strokewidth specification.

If the character sizes are intermixed, the line separation limitation for any pair of lines shall be that applicable to the largest character in the two lines.

**3.10. Character Skew.** The skew of a character is the rotational deviation of the printed image from its intended orientation relative to a document reference edge. Character skew shall not exceed 3 degrees.



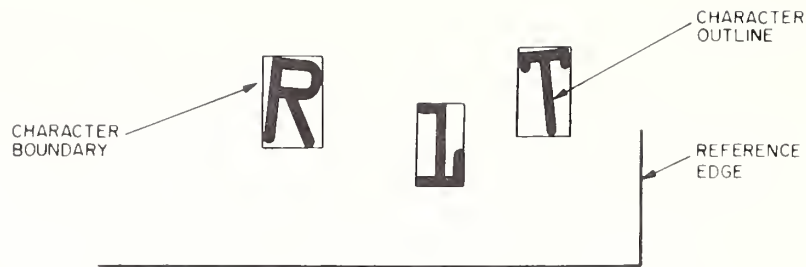


Figure II-4  
Character Boundaries

**3.11. Character Boundary.** The character boundary (see Figure II-4) is defined as the rectangle with sides parallel and perpendicular to a document reference edge which is drawn tangential to the character outline and contains the character completely. Skewed characters still have boundaries parallel or perpendicular to a form reference edge.

For the purpose of determining the boundary of the Long Vertical Mark, only that portion of the Long Vertical Mark which lies between the extension of the uppermost and lowermost horizontal boundaries of the adjacent character(s) will be considered.

The character boundary is used to measure character and line separation and to determine line boundary.

**3.12. Character Reference Lines.** Character reference lines (see Figures II-4 and II-5) are used to determine the position of a character relative to some other character or to some reference edge.

**3.12.1. Character Base Line.** The character base line is a reference line used to specify the nominal relative vertical position of a character relative to the line of type. The position of the base line is indicated on the drawings of all characters. (See Figures II-11 through II-96).

All characters should be printed with their base lines as close as practical to a common line. Deviations from character position with respect to the base line are permitted to achieve improved appearance as long as character misalignment (Section 3.15) specifications are not exceeded. For example, punctuations such as comma and semicolon may be positioned below the base line to achieve more conventional appearance, particularly when used with lower case alphabets. The Long Vertical Mark (Figure II-24) has no nominal vertical position.

**3.12.2. Average Base Line.** The average base line for a line or line segment is a horizontal line parallel or perpendicular to a reference edge, which passes through the average of the individual base line of all the characters in that line or line segment (see Figure II-3).

**3.12.3. Displacement from Base Line.** The base line displacement,  $Y$ , is the shortest distance between the base line and the lowest point on the nominal centerline of the character. The value of  $Y$  is zero for characters in which this point lies on the base line. For all the other characters, the value of  $Y$  is indicated on the corresponding character drawings.

**3.12.4. Character Spacing Reference Line.** The character spacing reference line is normally the vertical centerline of the character boundary, except on characters 4, f, and j which require a correction by the value  $\Delta X$ , as indicated on the corresponding character drawings.



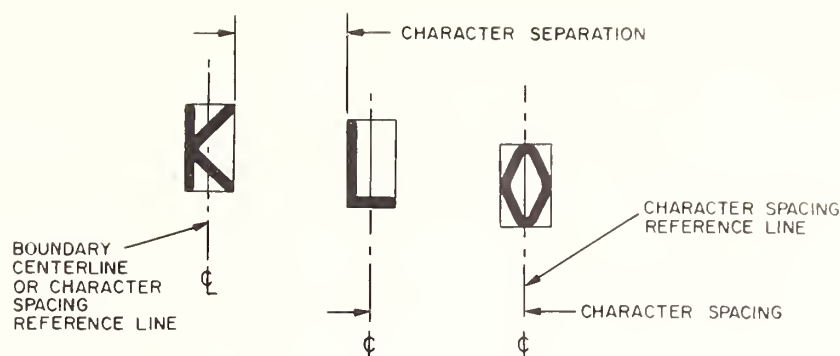


Figure II-5

## Character Separation and Spacing

3.13. **Character Spacing** (See Figure II-5). Character spacing is the horizontal distance between the character spacing reference lines of two adjacent characters (including the Long Vertical Mark). If one or both of the characters is either a 4, f, or j, this distance should be corrected by  $\Delta X$ , as shown in the corresponding character drawings.

Two characters are adjacent if the distance between their character spacing reference lines is smaller than the following maximum values:

Size	I	III	IV
Maximum Spacing.....	0.180 in (4.57 mm)	0.180 in (4.57 mm)	0.260 in (6.60 mm)

The distance between the character spacing reference lines of two adjacent characters shall not be less than the following specified minimum values:

Size	I	III	IV
Minimum Spacing.....	0.090 in (2.29 mm)	0.090 in (2.29 mm)	0.130 in (3.30 mm)

*Note:* Some journal tape printers may not provide a full character space for printing of the period, when used as a decimal point. As a result, the character spacing requirements of this paragraph cannot be met. Some OCR readers can permit this exception as long as the character separation requirements of Paragraph 3.14 is satisfied. When considering the installation of an OCR system of this type, close liaison with printer and scanner manufacturers is advised.

3.14. **Character Separation** (See Figure II-5). Character separation is the horizontal distance between the adjacent boundaries of any OCR character(s) and/or the Long Vertical Mark (see Figure II-24). The character separation shall not be less than the nominal strokewidth as specified in Section 2.1.

3.15. **Character Misalignment**. Character misalignment is the vertical distance "R" between the character base lines of two characters on the same line. Where characters do not normally touch the base line, the character misalignment may be determined by measuring the distance perpendicular to the average base line between the lowest stroke centerlines and correcting by  $\Delta Y$ , the difference between the Y values shown in the character drawings. See Figure II-6 and character drawings II-11 thru II-96.

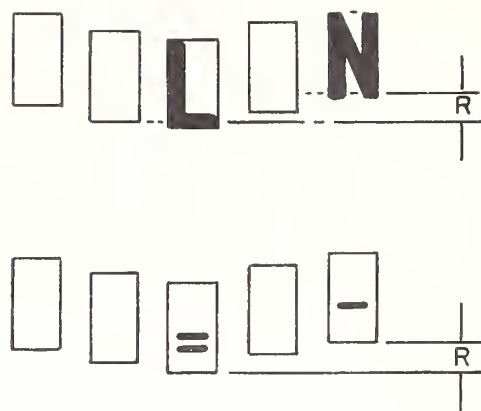


Figure II-6  
Character Misalignment

**3.15.1. Adjacent Character Misalignment.** Adjacent character misalignment is measured according to the above procedure. It shall not exceed the following values:

Size	I	III	IV
Max. adjacent character.....	0.027 in	0.035 in	0.042 in
Misalignment.....	(0.69 mm)	(0.89 mm)	(1.07 mm)

**3.15.2. Character Misalignment in a Line.** Character misalignment within a line is measured according to the above procedures. It shall not exceed:

Size	I	III	IV
Max. line character.....	0.054 in	0.070 in	0.085 in
Misalignment.....	(1.37 mm)	(1.78 mm)	(2.16 mm)

If more than one character size is used within a line or a line segment, such that the characters of different sizes are adjacent (or considered as part of the same data field), then the limitation applying to the smallest character size applies to the whole line or line segment.

**3.15.3. Long Vertical Mark Alignment.** The Long Vertical Mark must extend beyond the top and the bottom boundaries of any adjacent character (except when lower case characters with descenders are used). A Long Vertical Mark should not extend nearer than 0.1 in (2.5 mm) to an adjacent line boundary to which it does not apply.

TABLE II-2—Summary of Character Positioning Specifications—Style A

Size	Height		Min line spacing		Min line separation		Max adjacent char spacing		Min char spacing		Max adjacent misalignment		Max line misalignment	
	<i>in</i>	<i>mm</i>	<i>in</i>	<i>mm</i>	<i>in</i>	<i>mm</i>	<i>in</i>	<i>mm</i>	<i>in</i>	<i>mm</i>	<i>in</i>	<i>mm</i>	<i>in</i>	<i>mm</i>
I .....	0.094	2.39	0.157	3.99	0.025	0.64	0.180	4.57	0.090	2.29	0.027	0.69	0.054	1.37
III .....	0.126	3.20	0.188	4.78	0.060	1.52	0.180	4.57	0.090	2.29	0.035	0.89	0.070	1.78
IV .....	0.150	3.81	0.210	5.33	0.080	2.03	0.260	6.60	0.130	3.30	0.042	1.07	0.085	2.16

#### 4. Illustrative Character Drawings

Following are 1:1 and 5:1 representations of the standard character set in each of the three sizes specified, i.e., Size I, Size III and Size IV.\*

A B C D E F G H I J K L M  
 N O P Q R S T U V W X Y Z  
 a b c d e f g h i j k l m  
 n o p q r s t u v w x y z  
 0 1 2 3 4 5 6 7 8 9  
 . , : ; = + / \$ \* " & |  
 ' - { } % ? [ \ ^ \_ ` ~  
 ? . , ' -

Size I

A B C D E F G H I J K L M  
 N O P Q R S T U V W X Y Z  
 0 1 2 3 4 5 6 7 8 9  
 . , : ; = + / \$ \* " & |  
 { } % ? [ \ ^ \_  
 ? . , ' -

Size III

A B C D E F G H I J K L M  
 N O P Q R S T U V W X Y Z  
 0 1 2 3 4 5 6 7 8 9  
 . , : ; = + / \$ \* " & |  
 ' - { } % ? [ \ ^ \_ ` ~  
 Ü Ñ Ä Ø Ö Å Æ £ ¥ ¤ ¨

Size IV

Figure II-7

1:1 Illustration of Standard Character Set  
In the three standard sizes—Size I, Size III, and Size IV

\*These are illustrations only. The dimensions of the characters shown may not be in exact agreement with the specifications contained in the Standard.

A B C D E F G H I J K L M  
 N O P Q R S T U V W X Y Z  
 a b c d e f g h i j k l m  
 n o p q r s t u v w x y z  
 0 1 2 3 4 5 6 7 8 9  
 . , : ; = + / \$ \* ^ & |  
 ' - { } % ? [ \ ] \_  
 ? . , ' -

Figure II-8

5:1 Illustration of Standard Character Set (Size I)

A B C D E F G H I J K L M  
 N O P Q R S T U V W X Y Z  
 0 1 2 3 4 5 6 7 8 9  
 . , : ; = + / \$ \* ^ & |  
 { } % ? [ \ ]  
 . , ? ' -

Figure II-9

5:1 Illustration of Standard Character Set (Size III)

A B C D E F G H I J K L M  
 N O P Q R S T U V W X Y Z  
 0 1 2 3 4 5 6 7 8 9  
 . , : ; = + / \$ \* ^ & |  
 ' - { } % ? [ \ ]  
 ' -  
 . , ?

Figure II-10

5:1 Illustration of Standard Character Set (Size IV)

TABLE II-3—OCR Character Dimension Equivalents

## Inch and Metric Measurement

		Inches			Millimeters		
		I	III	IV	I	III	IV
	<i>H</i>	0.0940	0.1260	0.1500	2.388	3.200	3.810
	<i>W</i>	.0550	.0600	.0800	1.397	1.524	2.032
	<i>T</i>	.0140	.0150	.0200	0.356	0.381	0.508
1/2	<i>T</i>	.0070	.0075	.0100	.178	.191	.254
3/2	<i>T</i>	.0210	.0225	.0300	.533	.572	.762
	<i>T</i>	.0280	.0300	.0400	.711	.762	1.016
1/8	<i>W</i>	.0069	.0075	.0100	.175	.190	0.254
1/4	<i>W</i>	.0138	.0150	.0200	.351	.381	.508
3/8	<i>W</i>	.0206	.0225	.0300	.523	.572	.762
1/2	<i>W</i>	.0275	.0300	.0400	.698	.762	1.016
5/8	<i>W</i>	.0344	.0375	.0500	.874	.952	1.270
3/4	<i>W</i>	.0413	.0450	.0600	1.050	1.143	1.524
1/16	<i>H</i>	.0059	.0079	.0094	0.150	0.201	0.239
1/8	<i>H</i>	.0118	.0158	.0188	.300	.401	.478
1/6	<i>H</i>	.016	—	—	.41	—	—
3/16	<i>H</i>	.0176	.0236	.0281	.447	.599	.714
1/4	<i>H</i>	.0235	.0315	.0375	.597	.800	.952
5/16	<i>H</i>	.0294	.0394	.0469	.747	1.001	1.191
3/8	<i>H</i>	.0353	.0473	.0563	.897	1.201	1.430
7/16	<i>H</i>	.0411	.0551	.0656	1.044	1.400	1.666
1/2	<i>H</i>	.0470	.0630	.0750	1.194	1.600	1.905
9/16	<i>H</i>	.0529	.0709	.0844	1.344	1.801	2.144
5/8	<i>H</i>	.0588	.0788	.0938	1.494	2.002	2.363
11/16	<i>H</i>	.0646	.0866	.1031	1.641	2.200	2.613
3/4	<i>H</i>	.0705	.0945	.1125	1.791	2.400	2.858
13/16	<i>H</i>	.0764	.1024	.1219	1.941	2.601	3.096
5/6	<i>H</i>	.078	—	—	1.98	—	—
7/8	<i>H</i>	.0823	.1103	.1313	2.090	2.802	3.335
15/16	<i>H</i>	.0881	.1181	.1406	2.238	3.000	3.571
	<i>r</i> <sub>1</sub>	.0248	.0401	.0431	0.630	1.019	1.095
	<i>r</i> <sub>2</sub>	.0111	.0112	.0156	.282	0.285	0.396
	<i>r</i> <sub>3</sub>	.0100	.0105	.0143	.254	.267	.363
	<i>r</i> <sub>4</sub>	.0087	.0172	.0142	.221	.310	.361
	<i>L</i>	.146	.196	.233	3.71	4.98	5.91

Formulas for Radii *r*<sub>1</sub> · *r*<sub>4</sub>

$$r_1 = \frac{H}{16} \left( \frac{7H}{6W} + \sqrt{\left( \frac{7H}{6W} \right)^2 + 1} \right) \quad r_3 = \frac{W}{8} \left( 1 - \frac{H}{2W} + \sqrt{\left( \frac{H}{2W} \right)^2 + 1} \right)$$

$$r_2 = \frac{W}{8} \left( \frac{6W}{7H} + \sqrt{\left( \frac{6W}{7H} \right)^2 + 1} \right) \quad r_4 = \frac{H}{16} \left( 1 - \frac{4W}{3H} + \sqrt{\left( \frac{4W}{3H} \right)^2 + 1} \right)$$



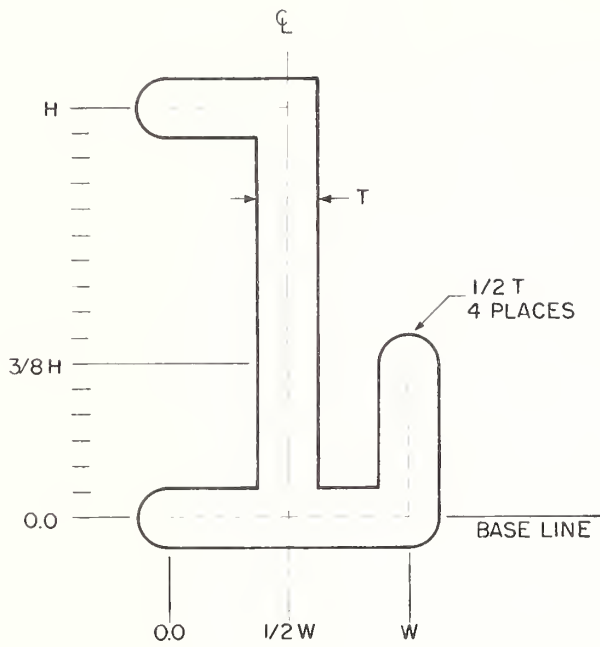


Figure II-11  
Number One

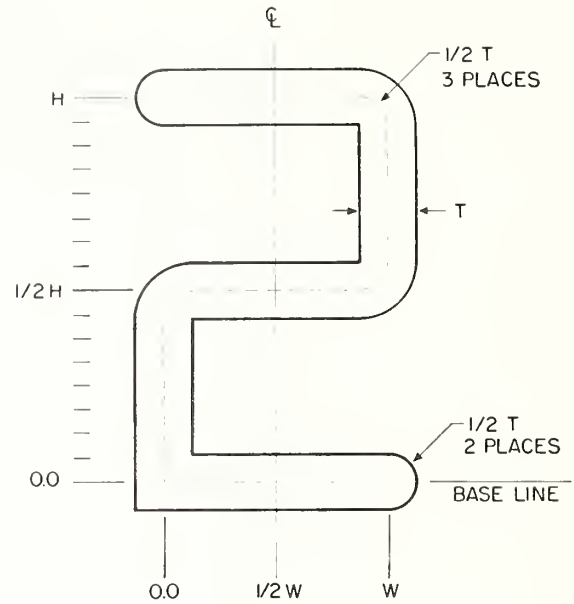


Figure II-12  
Number Two

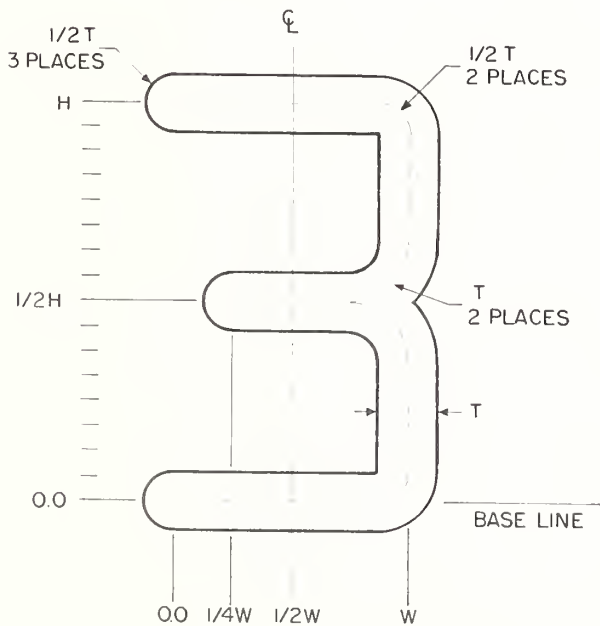


Figure II-13  
Number Three

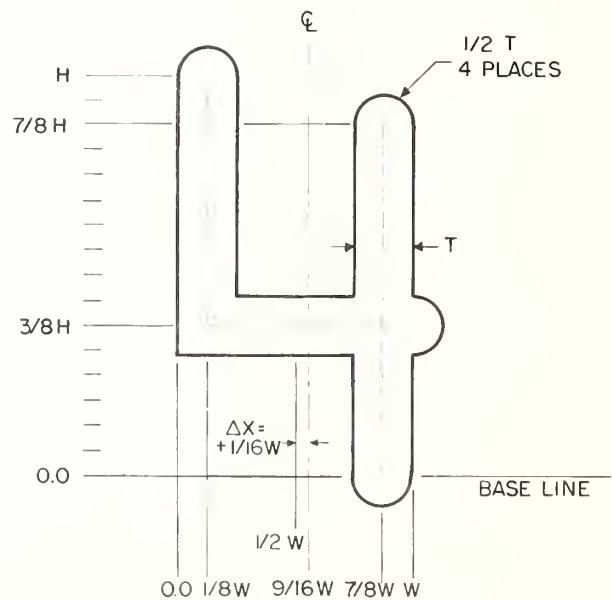


Figure II-14  
Number Four

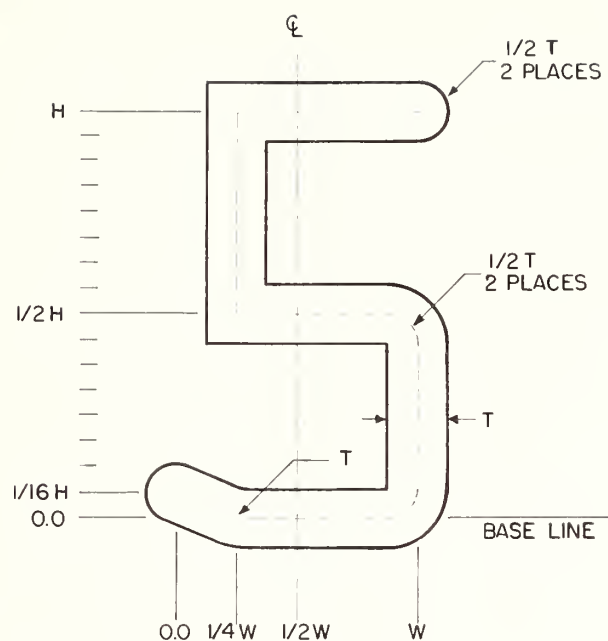


Figure II-15  
Number Five

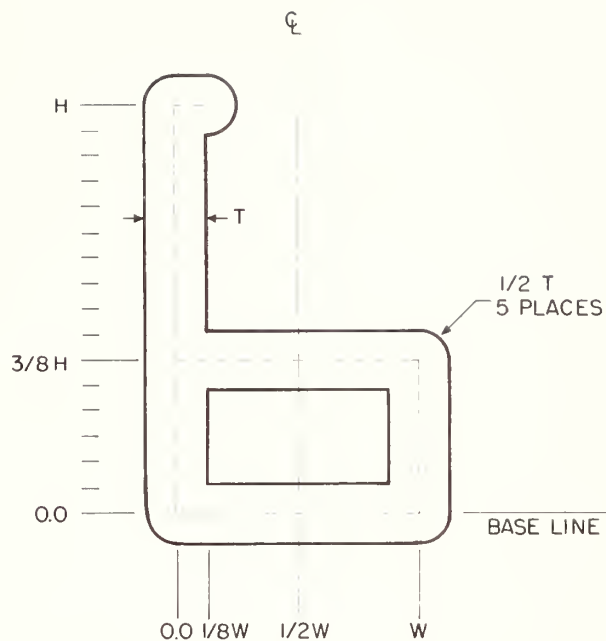


Figure II-16  
Number Six

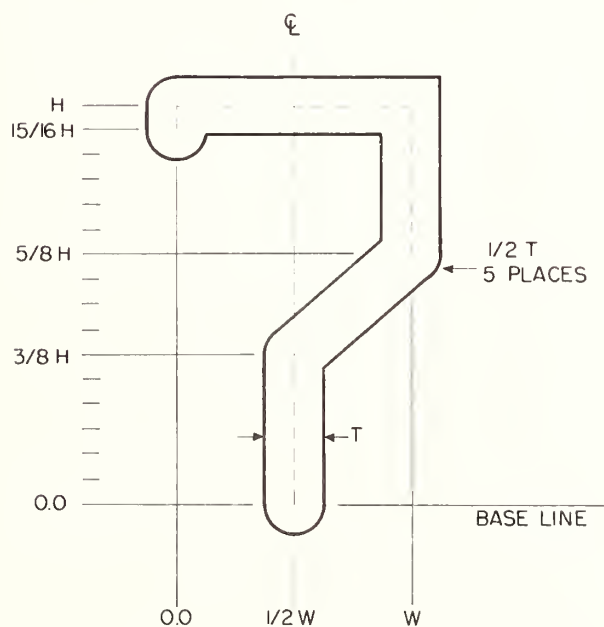


Figure II-17  
Number Seven

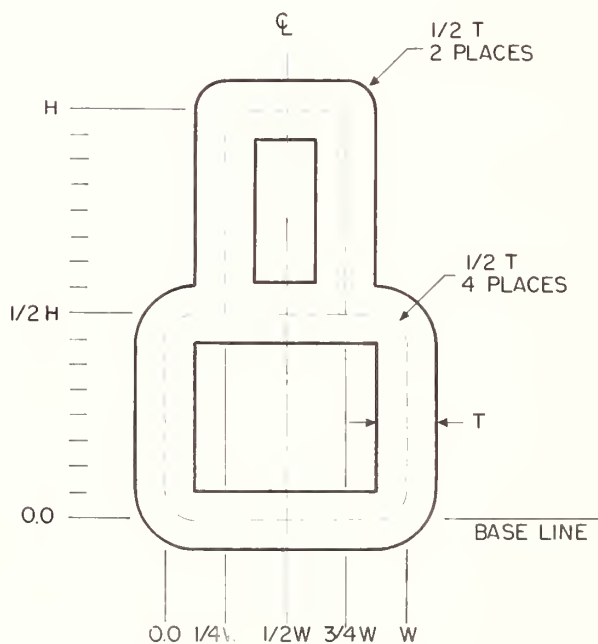


Figure II-18  
Number Eight

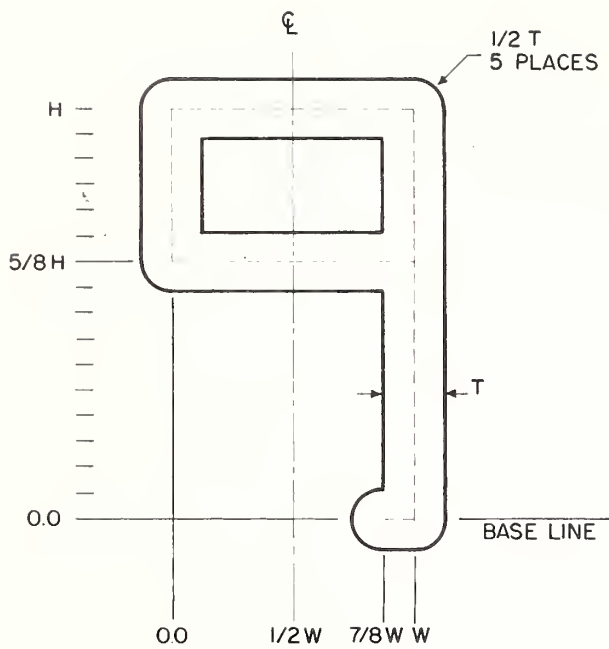


Figure II-19  
Number Nine

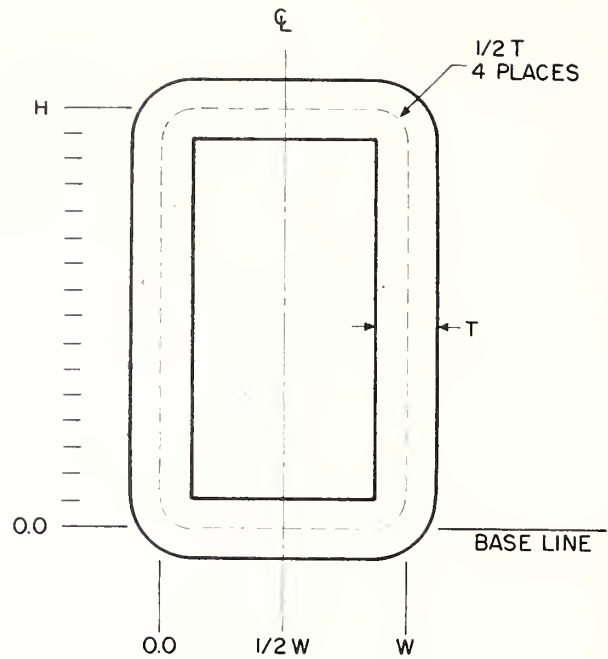


Figure II-20  
Number Zero

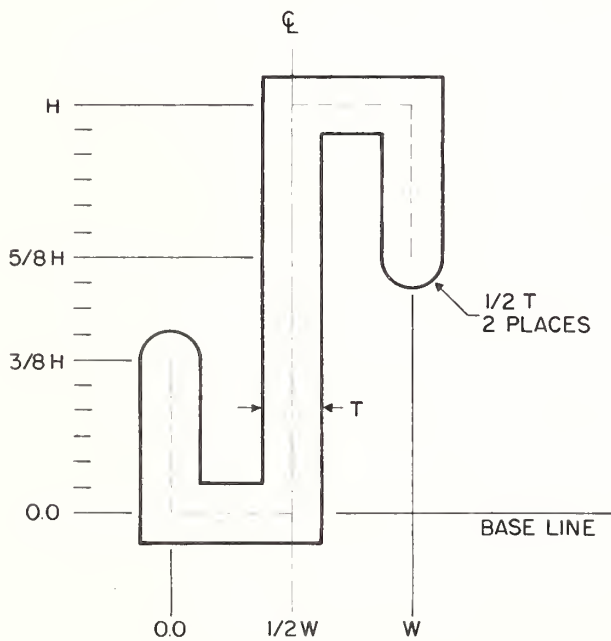


Figure II-21  
Symbol Hook

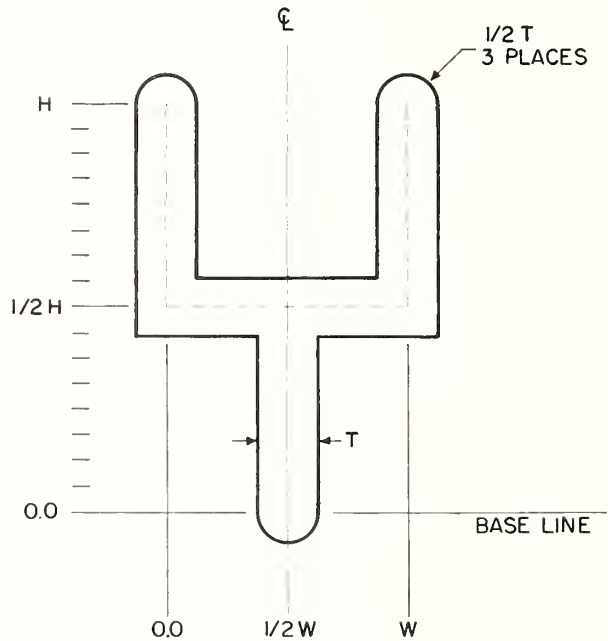


Figure II-22  
Symbol Fork

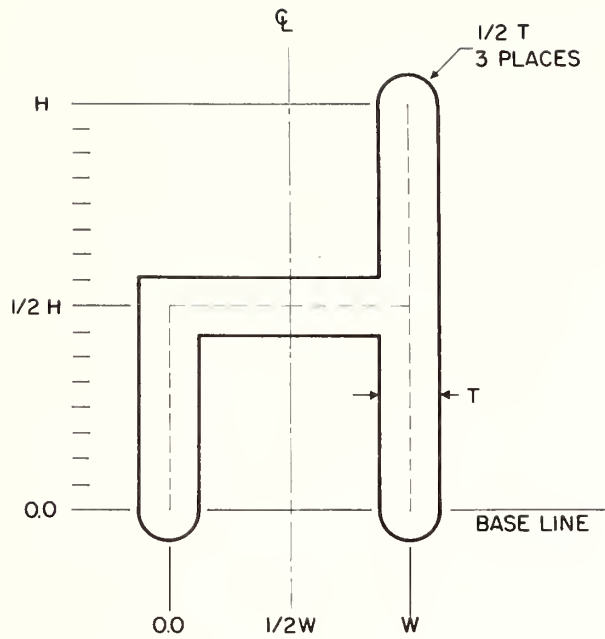


Figure II-23  
Symbol Chair

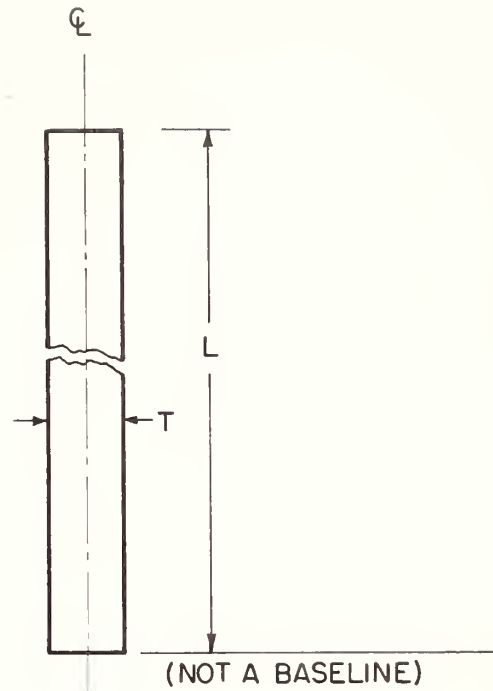


Figure II-24  
Long Vertical Mark

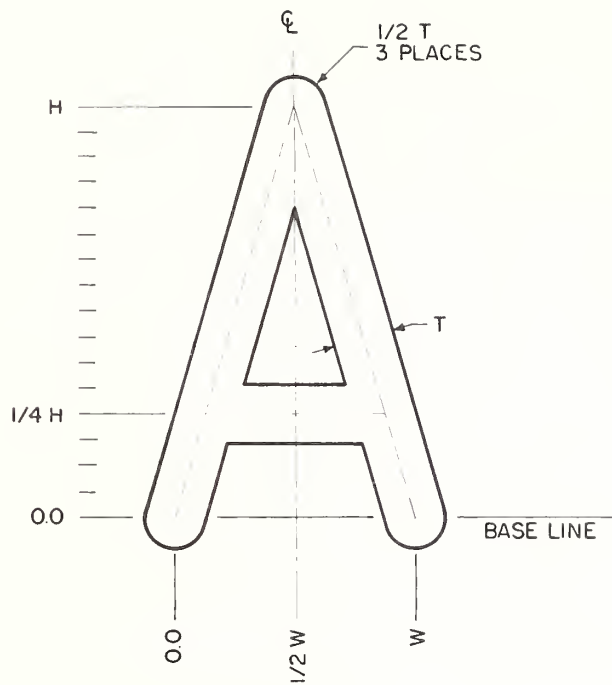


Figure II-25  
Letter A

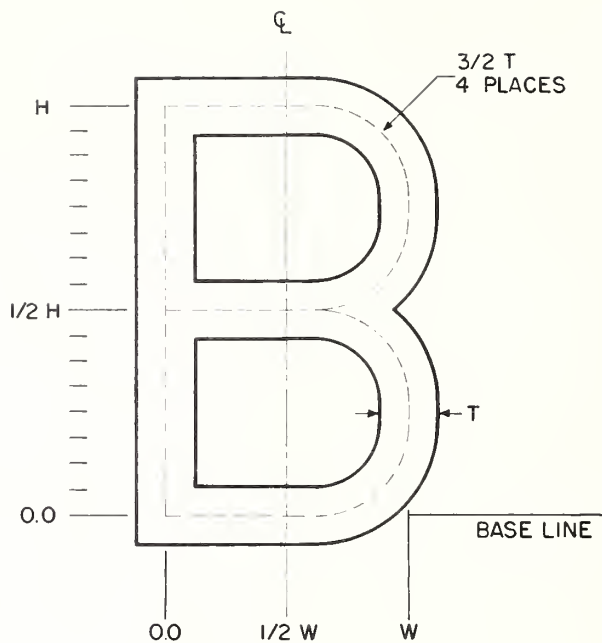


Figure II-26  
Letter B

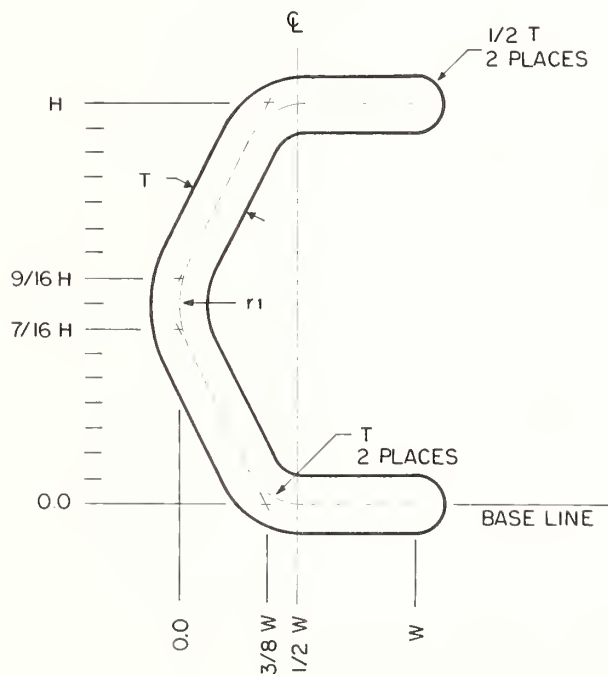


Figure II-27  
Letter C

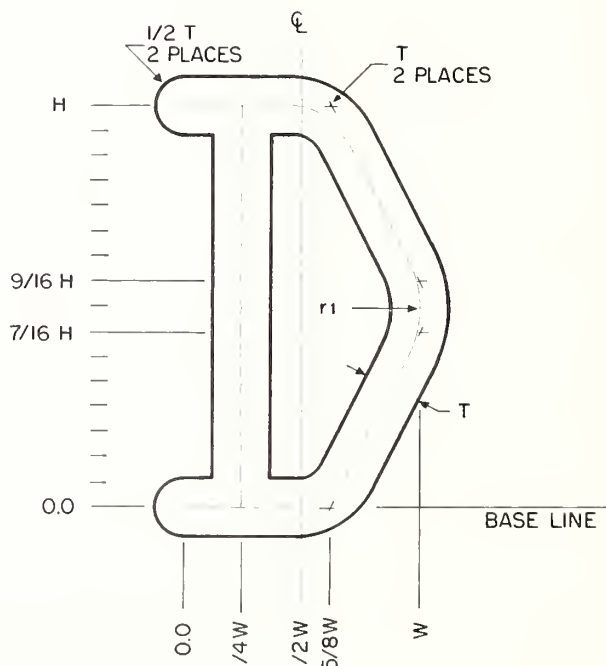


Figure II-28  
Letter D



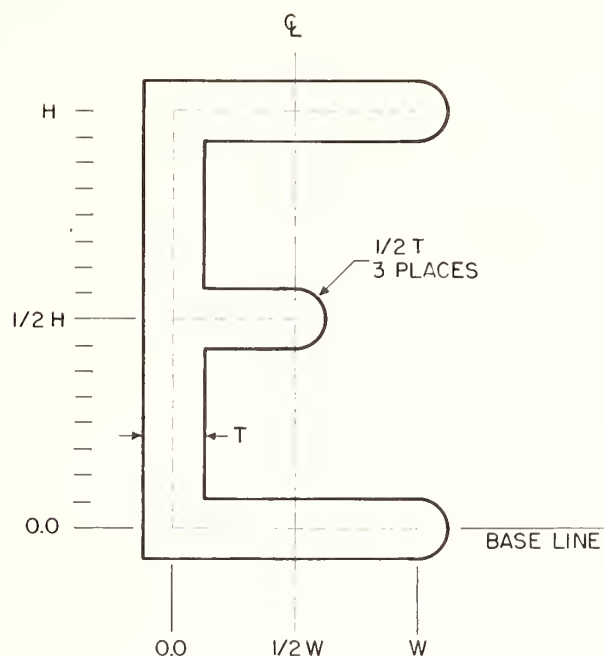


Figure II-29  
Letter E

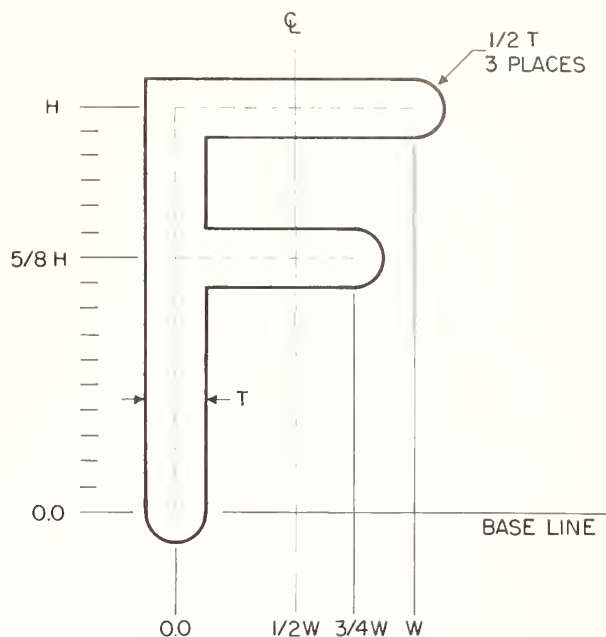


Figure II-30  
Letter F

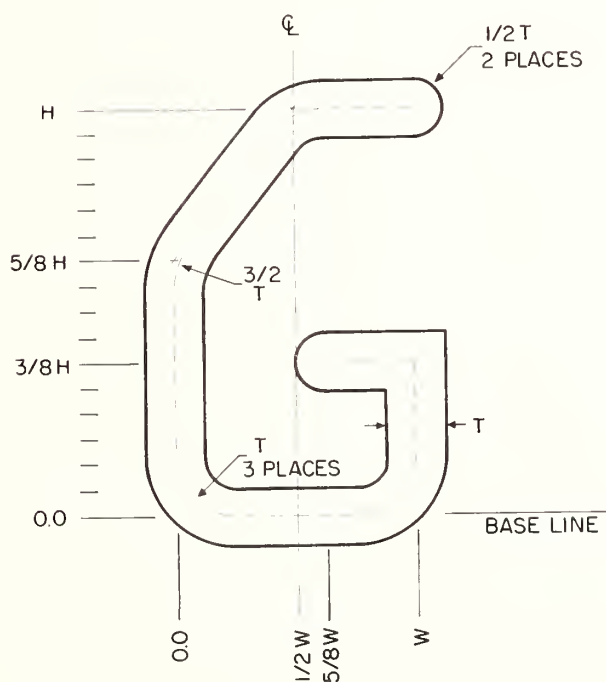


Figure II-31  
Letter G

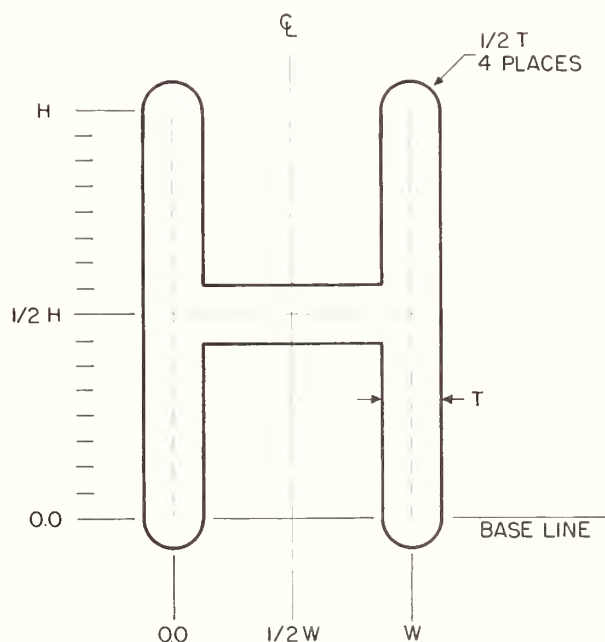


Figure II-32  
Letter H

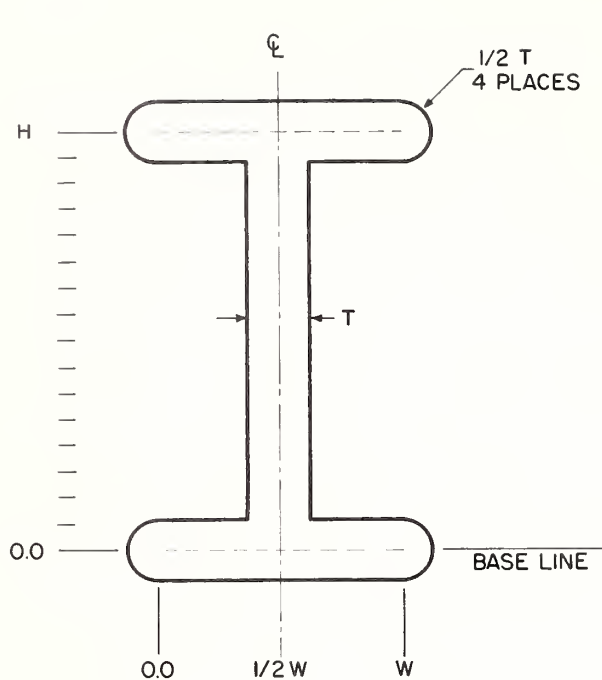


Figure II-33  
Letter I

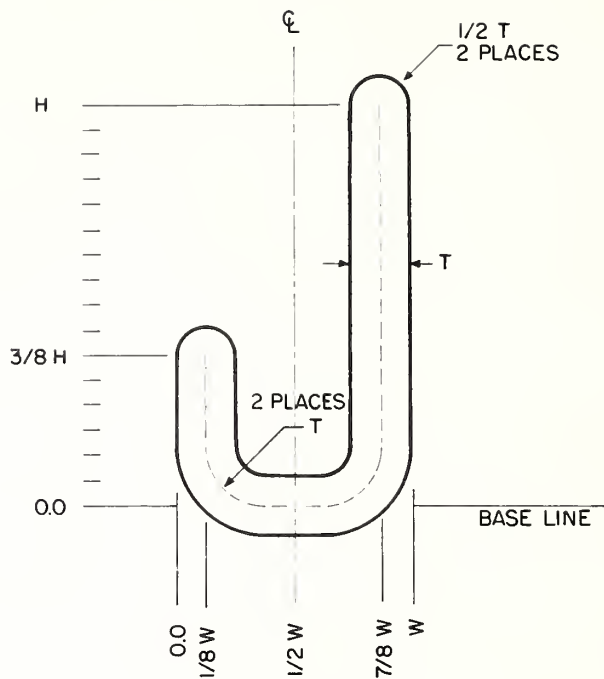


Figure II-34  
Letter J

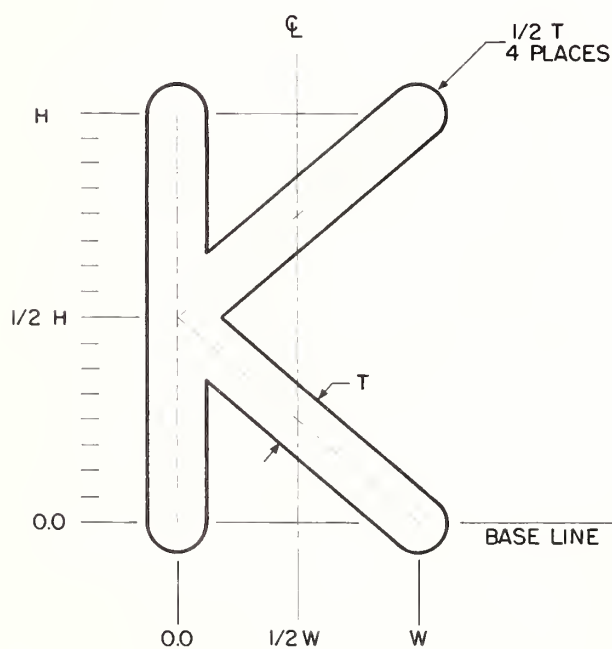


Figure II-35  
Letter K

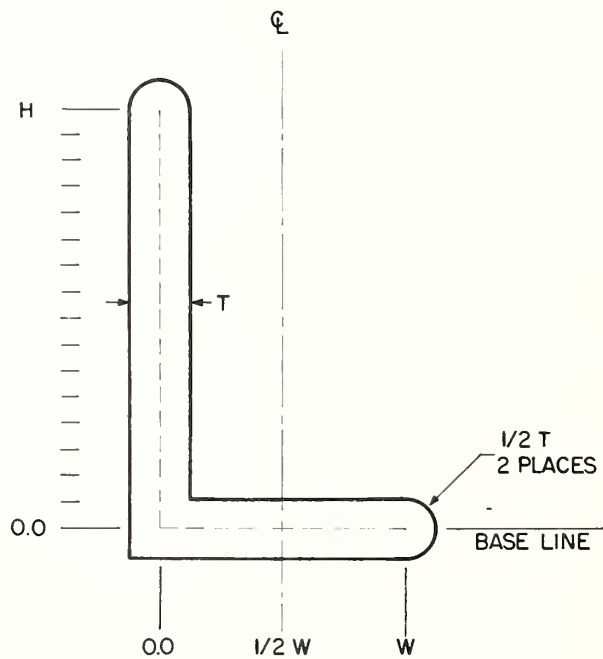


Figure II-36  
Letter L

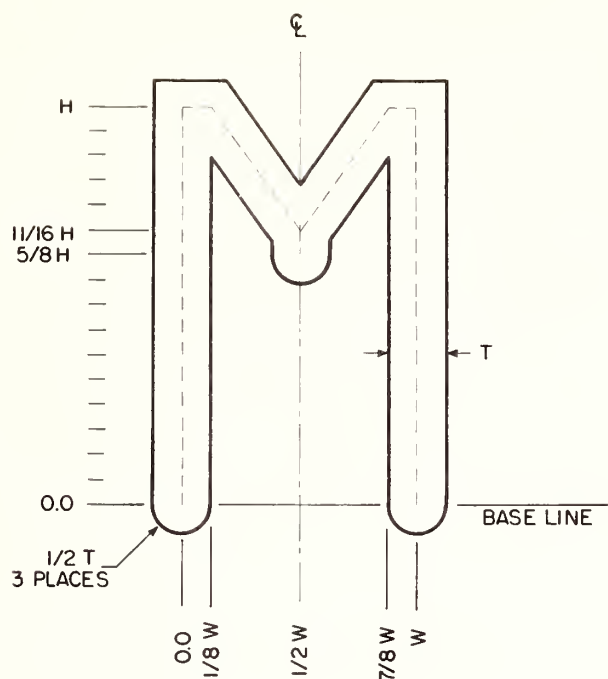


Figure II-37  
Letter M

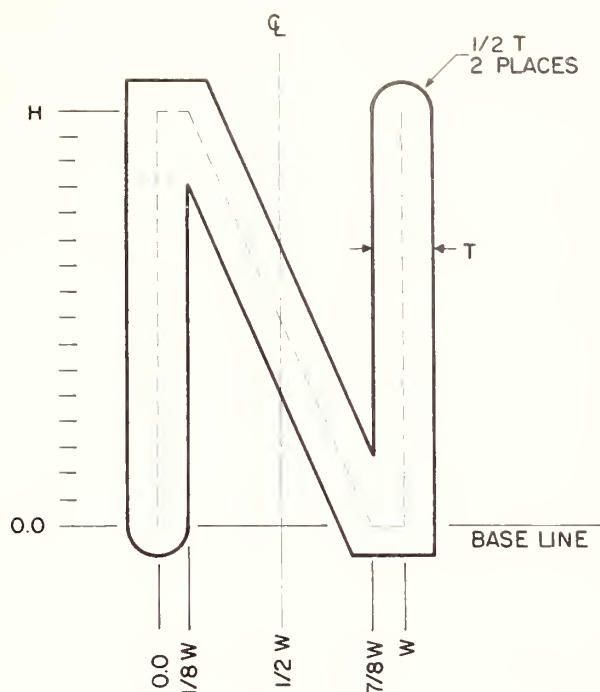


Figure II-38  
Letter N

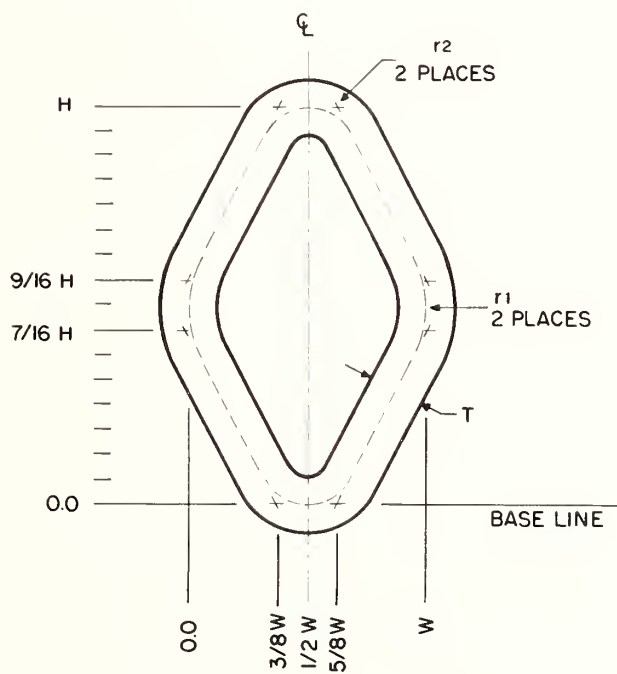


Figure II-39  
Letter O

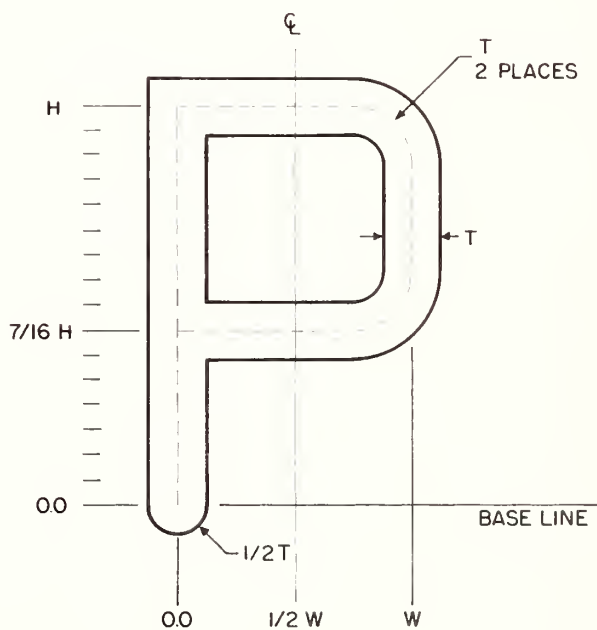


Figure II-40  
Letter P

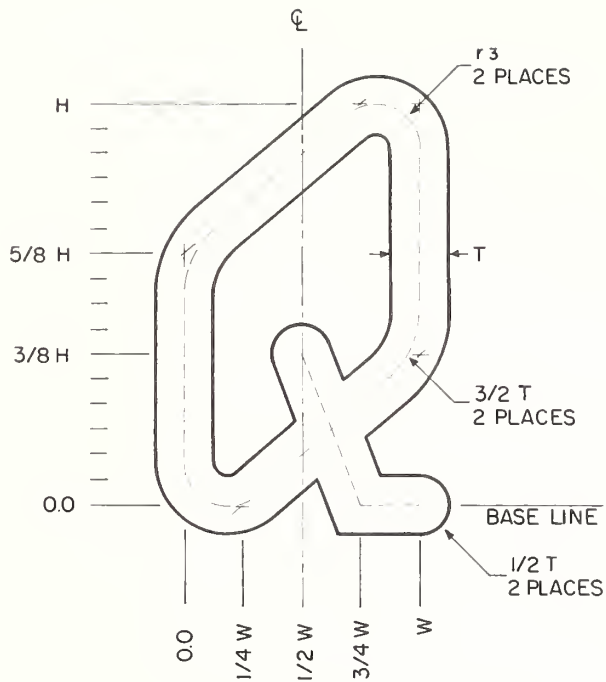


Figure II-41  
Letter Q

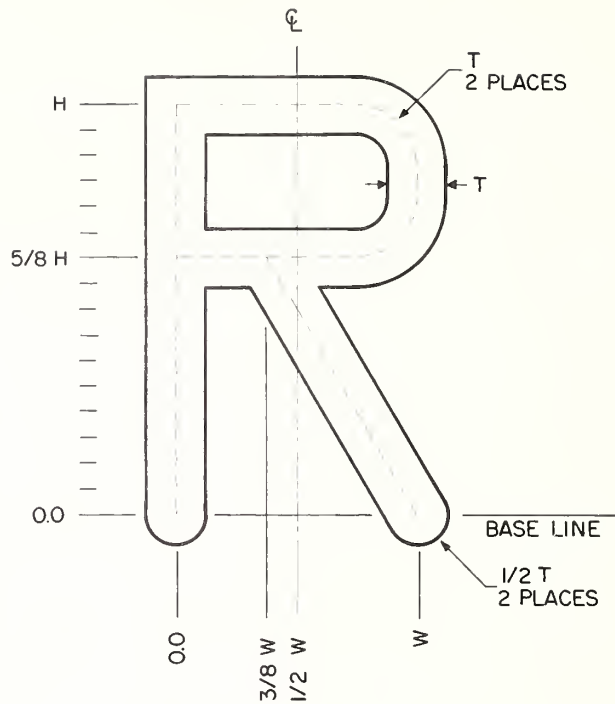


Figure II-42  
Letter R

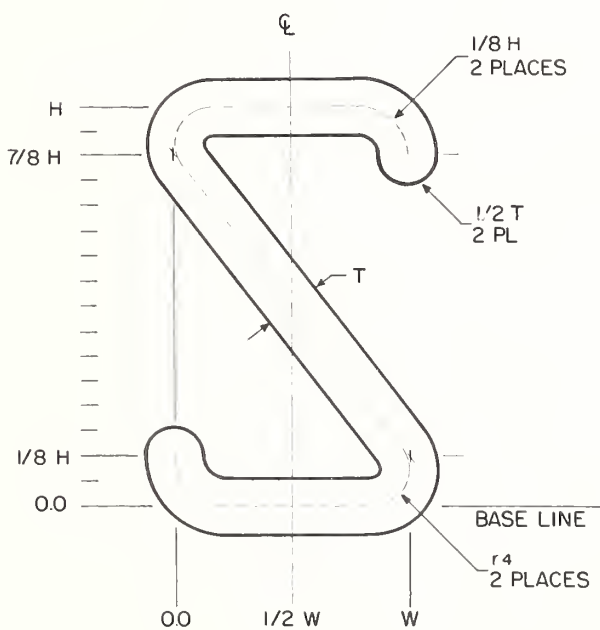


Figure II-43  
Letter S

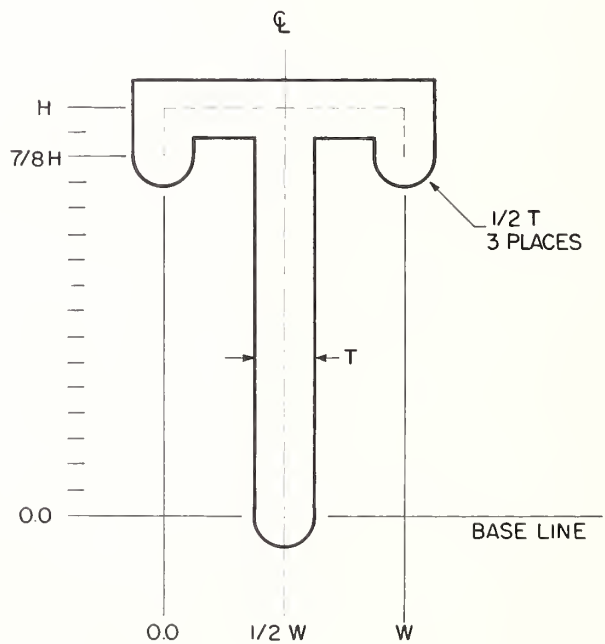


Figure II-44  
Letter T

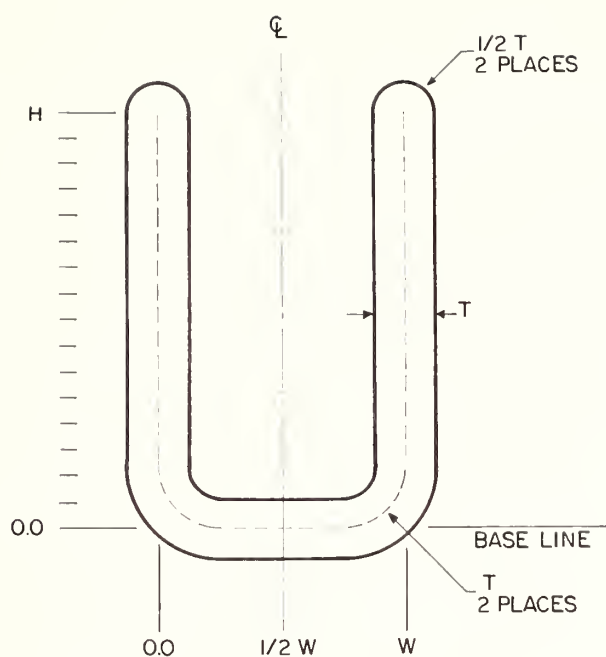


Figure II-45  
Letter U

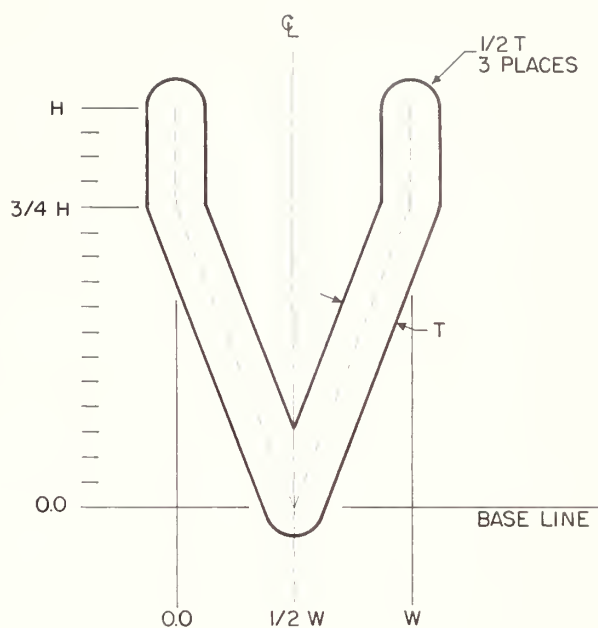


Figure II-46  
Letter V

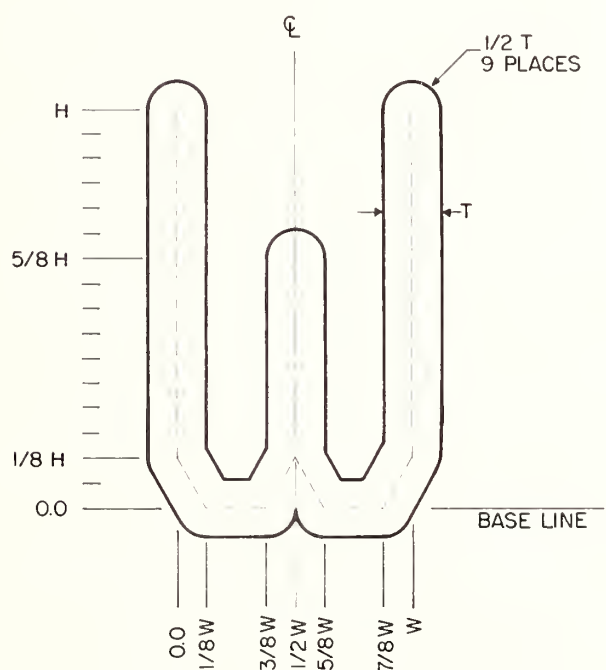


Figure II-47  
Letter W

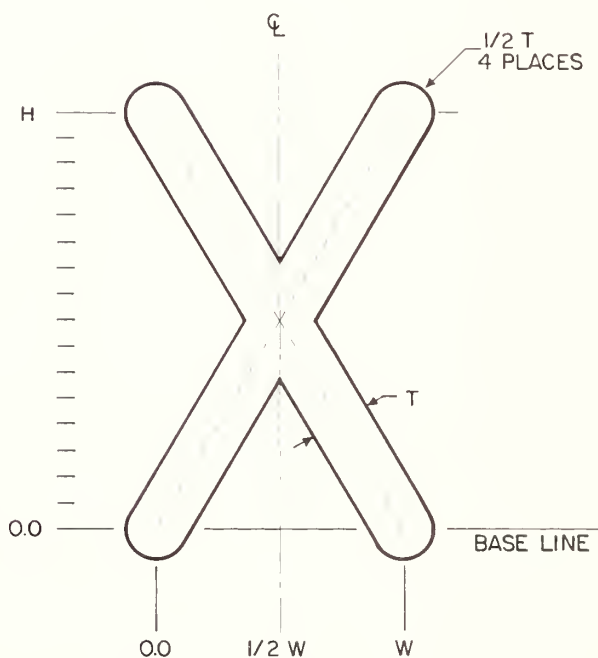


Figure II-48  
Letter X



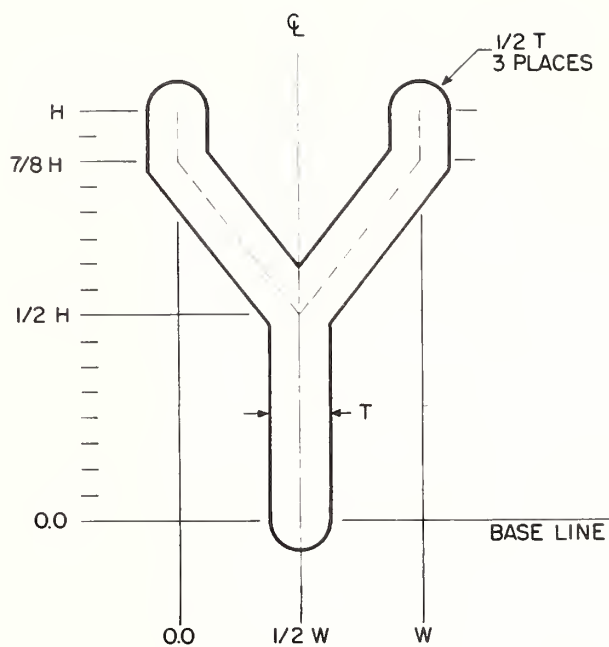


Figure II-49  
Letter Y

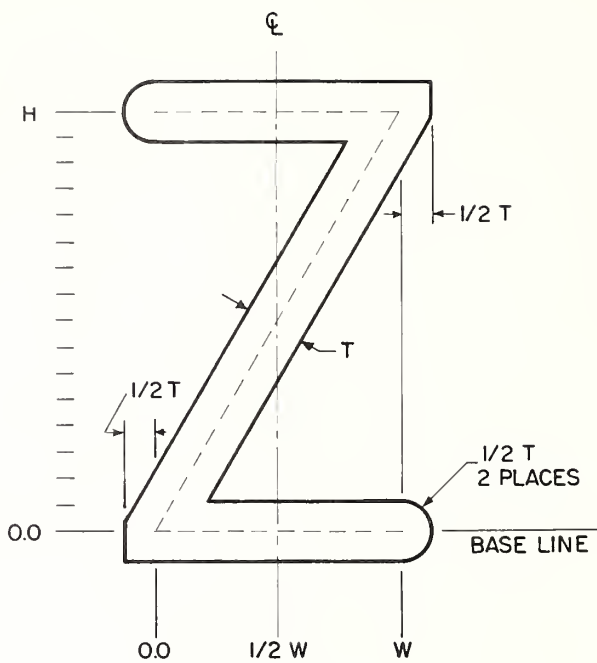


Figure II-50  
Letter Z

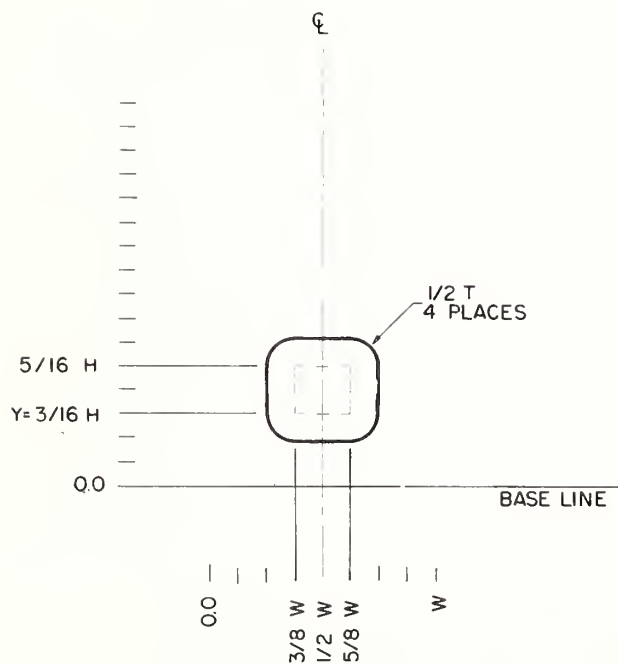


Figure II-51  
Period

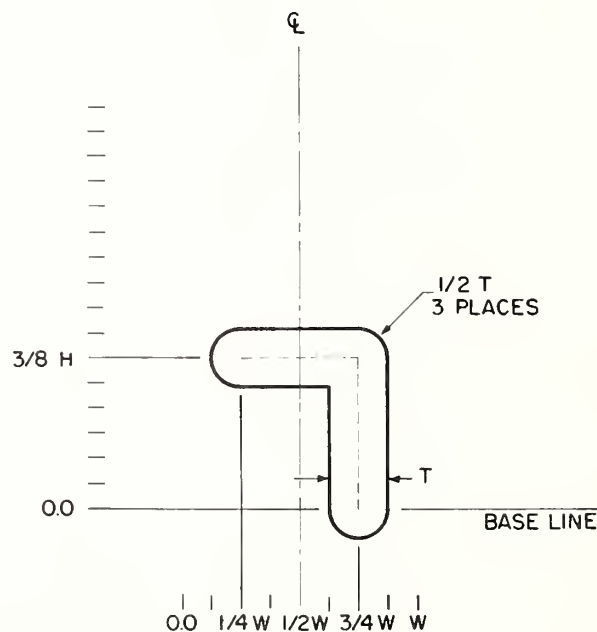


Figure II-52  
Comma

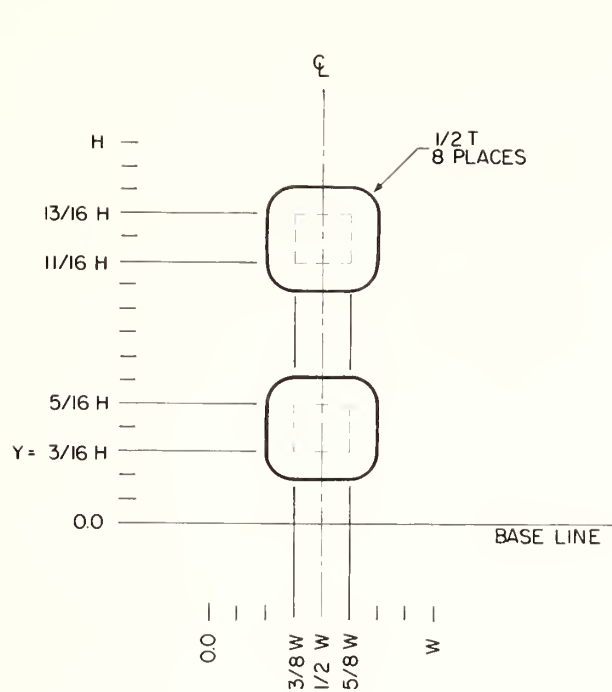


Figure II-53  
Colon

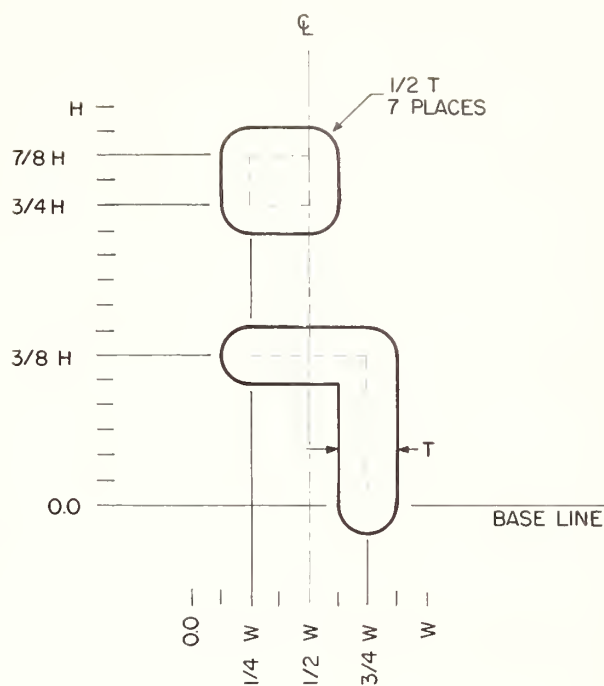


Figure II-54  
Semicolon

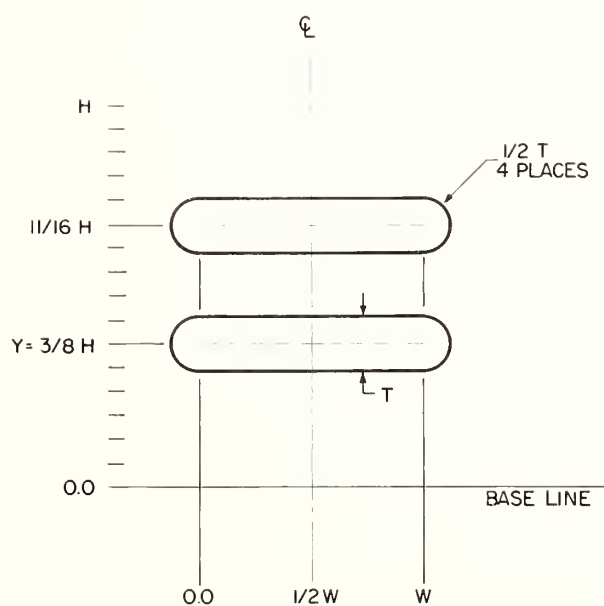


Figure II-55  
Equals

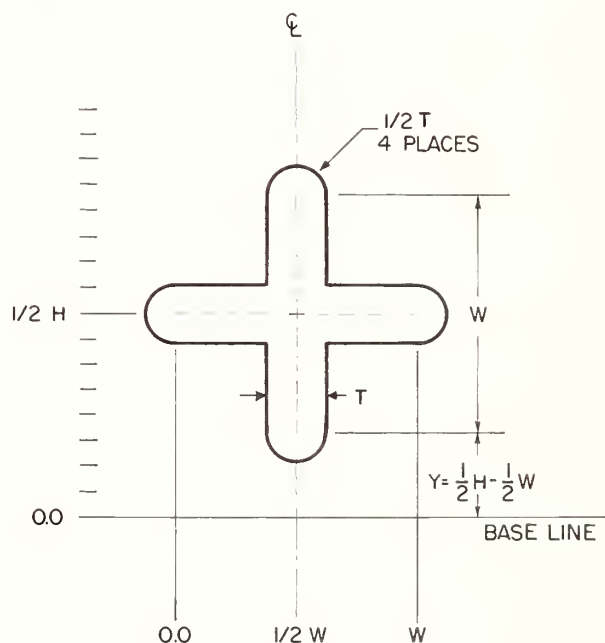


Figure II-56  
Plus

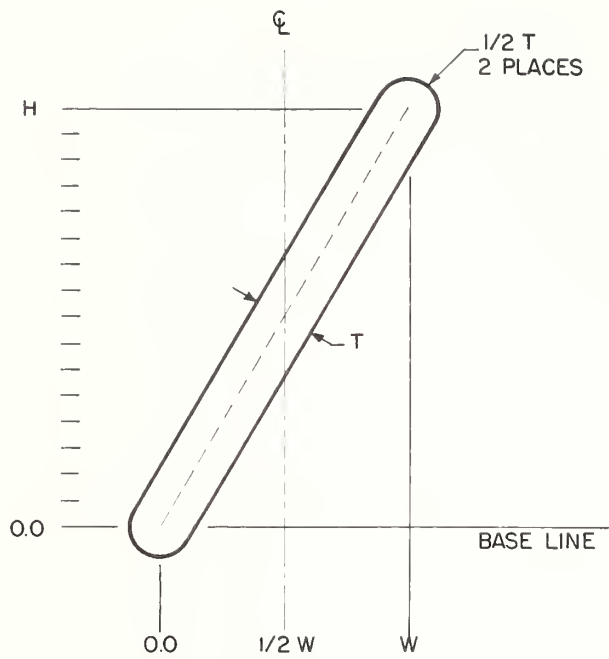


Figure II-57  
Slant

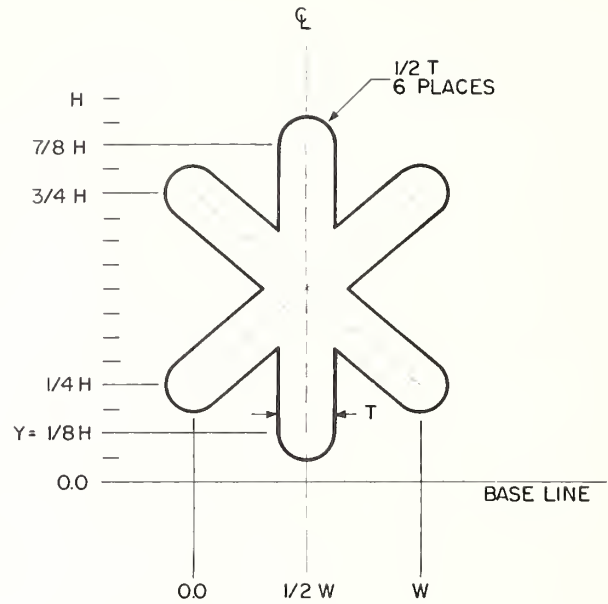


Figure II-58  
Asterisk

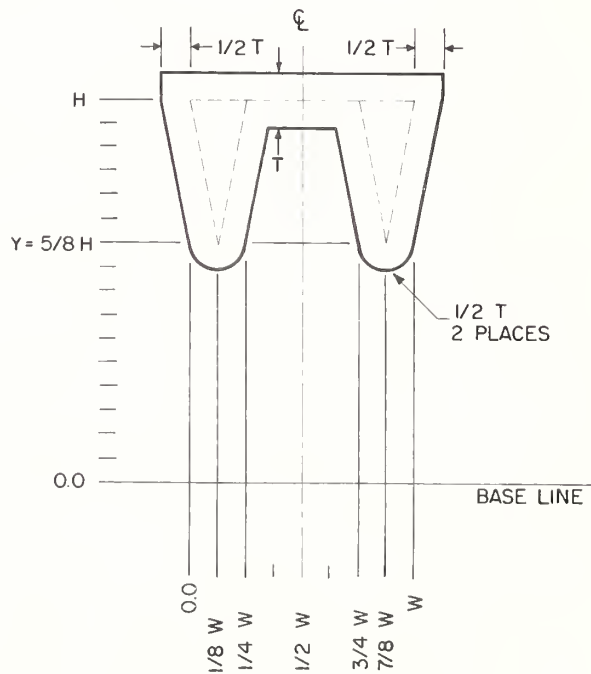


Figure II-59  
Quotation Mark

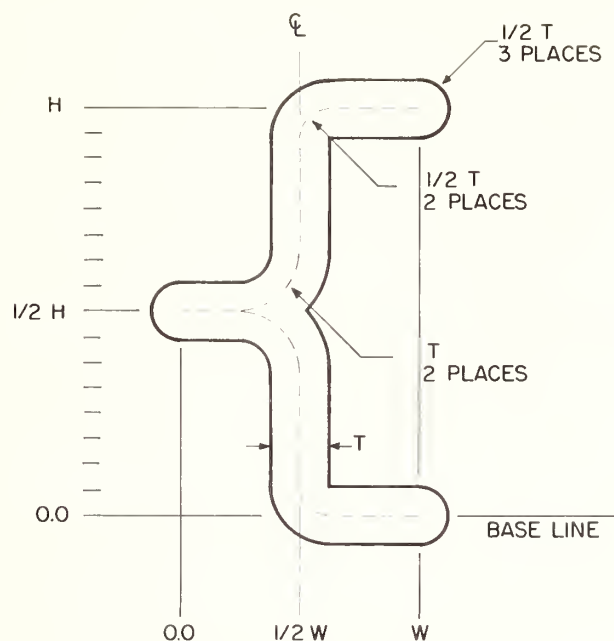


Figure II-60  
Opening Parenthesis

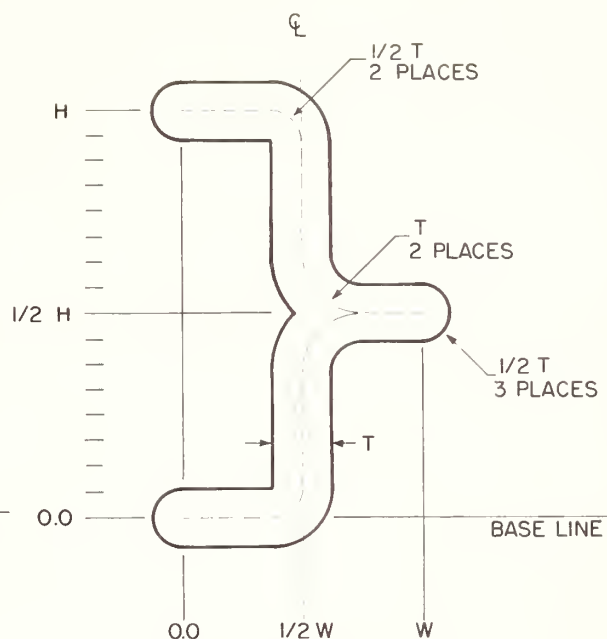


Figure II-61  
Closing Parenthesis

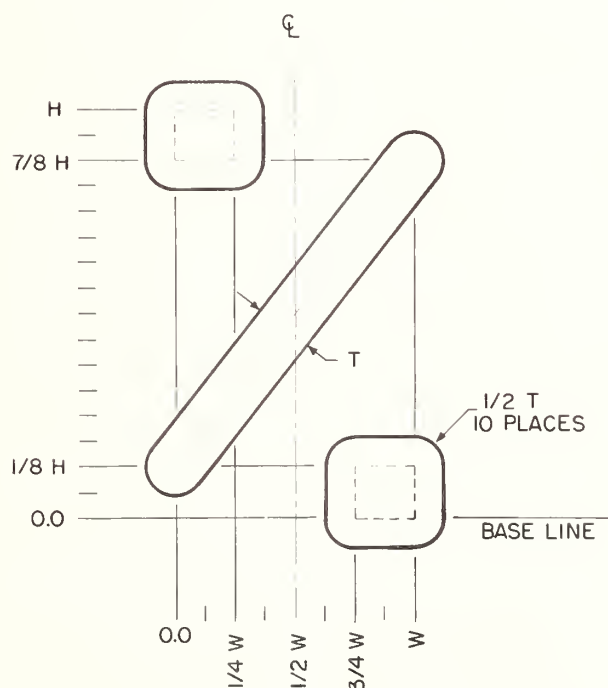


Figure II-62  
Percent

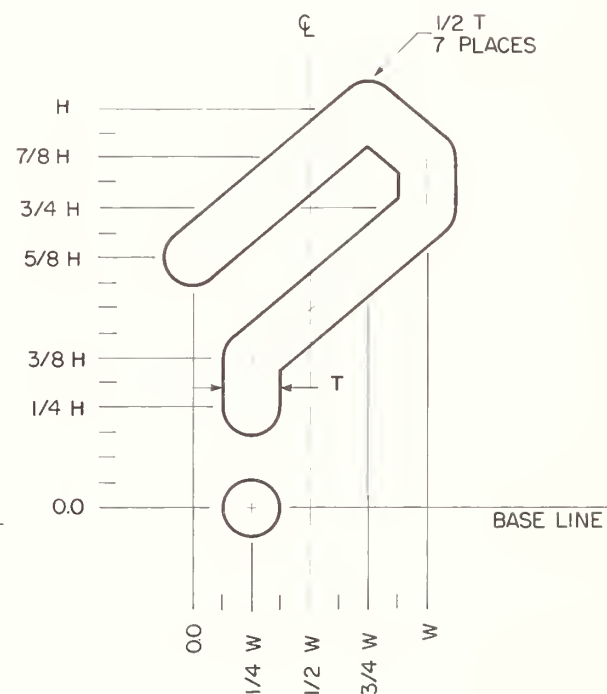


Figure II-63  
Question Mark

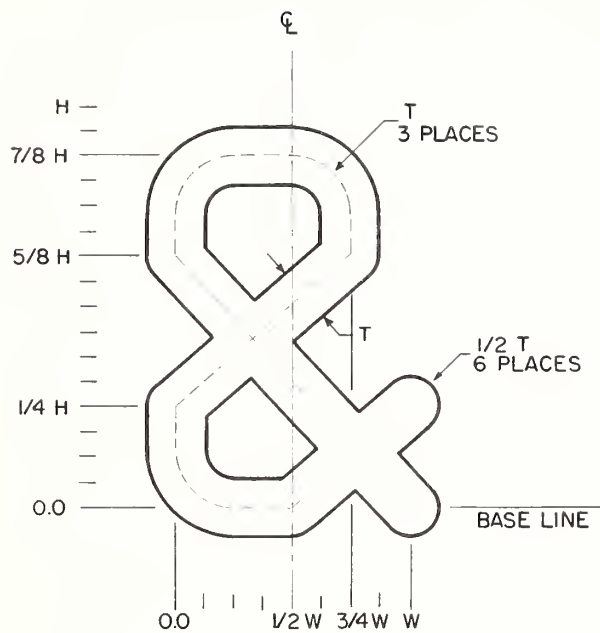


Figure II-64  
Ampersand

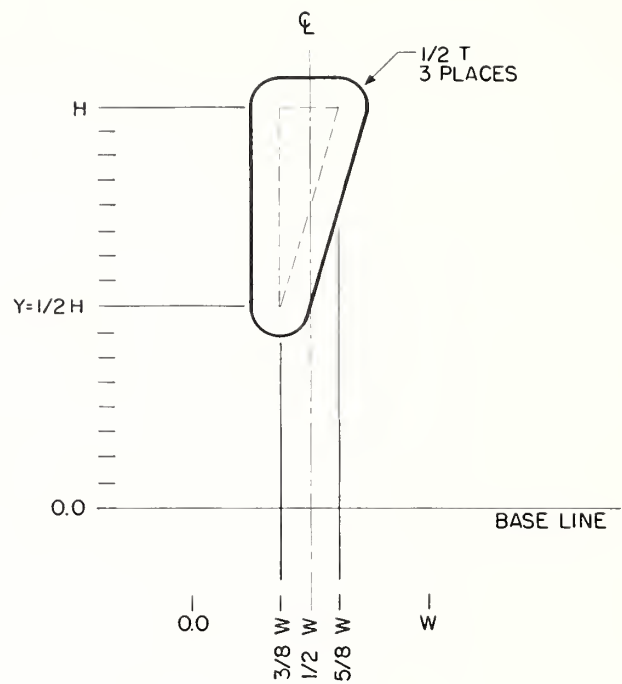


Figure II-65  
Apostrophe

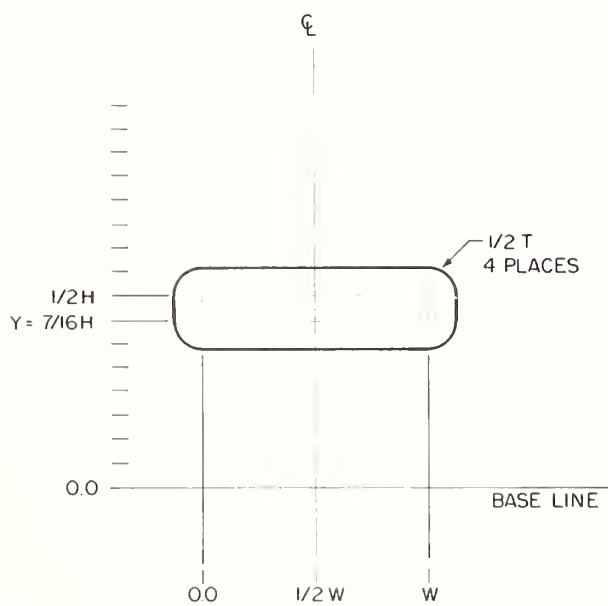


Figure II-66  
Hyphen

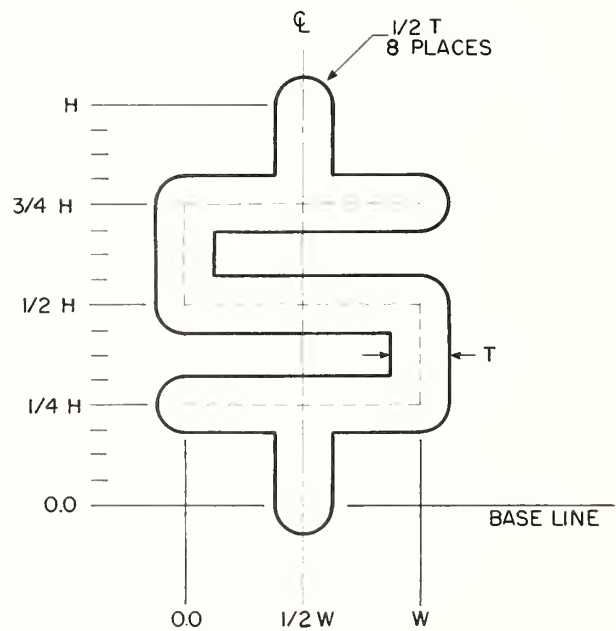


Figure II-67  
Dollar Sign



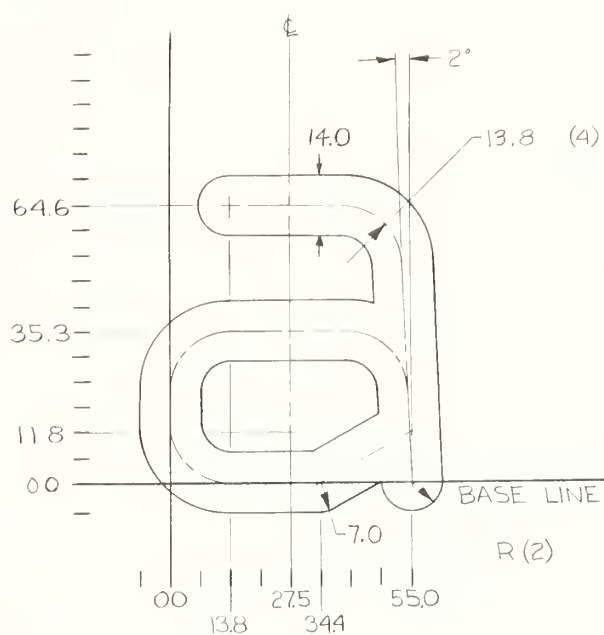


Figure II-68  
Small Letter A

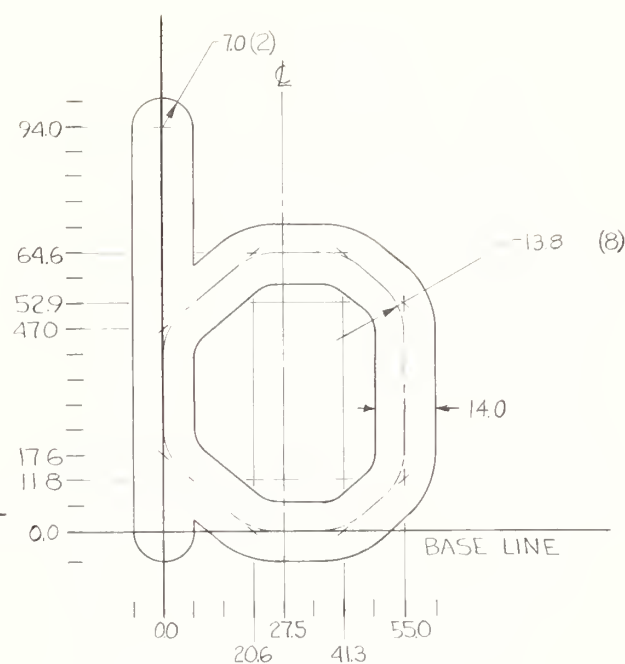


Figure II-69  
Small Letter B

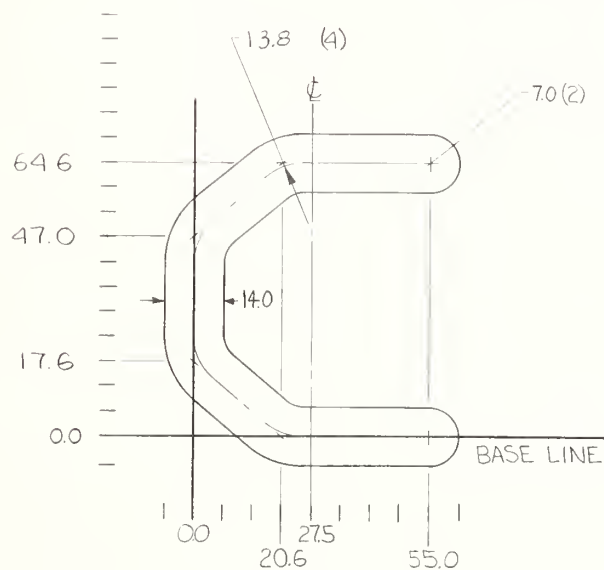


Figure II-70  
Small Letter C

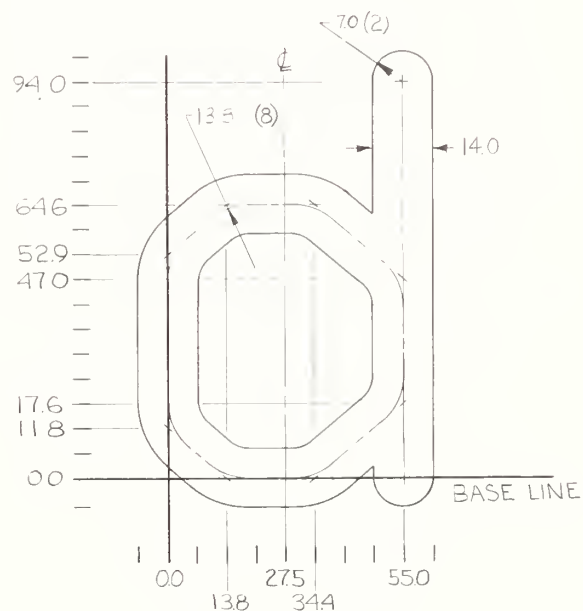


Figure II-71  
Small Letter D

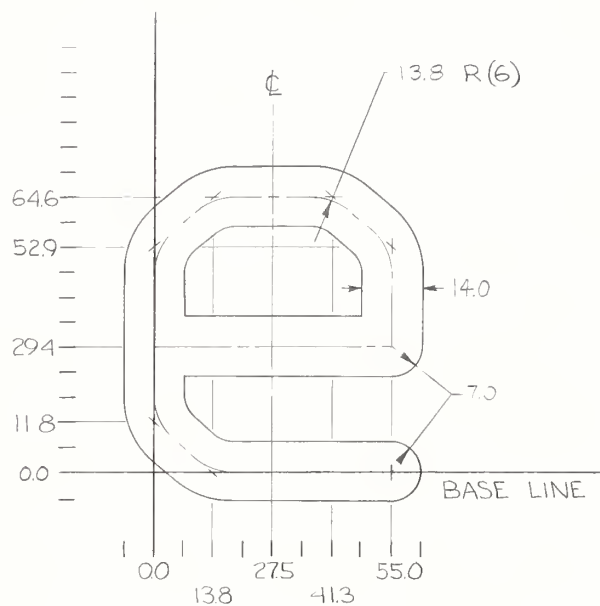


Figure II-72  
Small Letter E

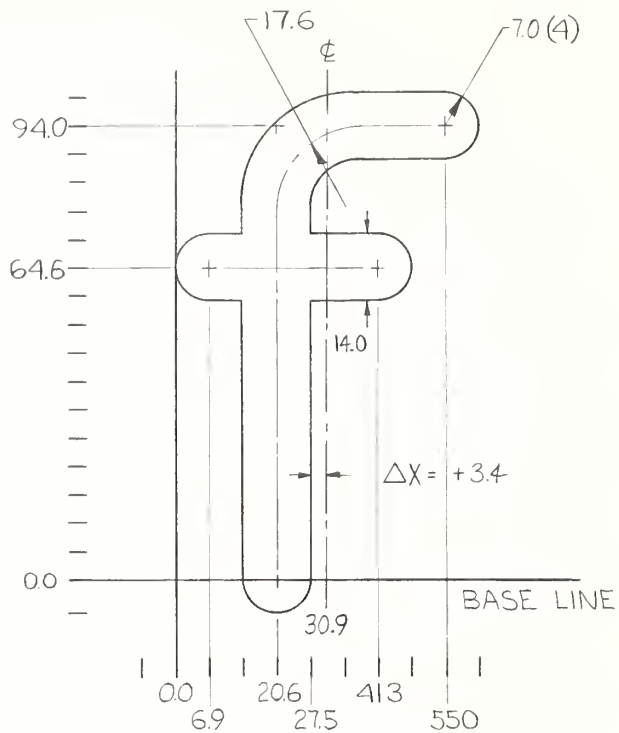


Figure II-73  
Small Letter F

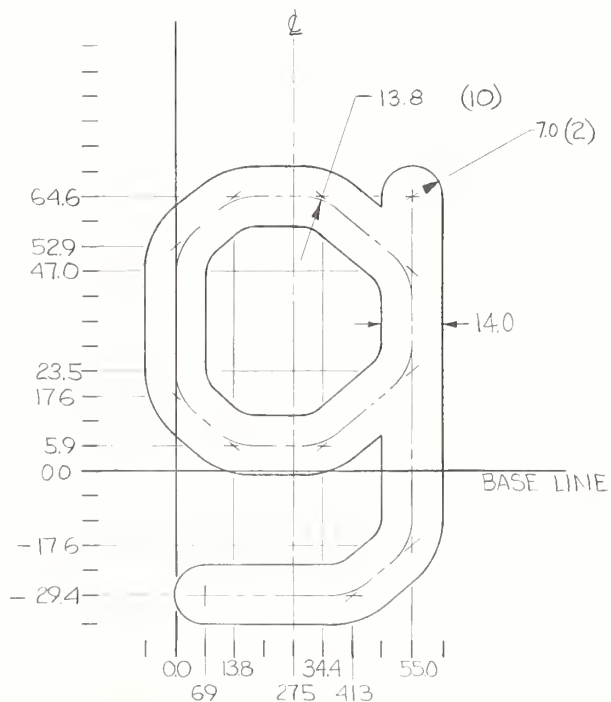


Figure II-74  
Small Letter G

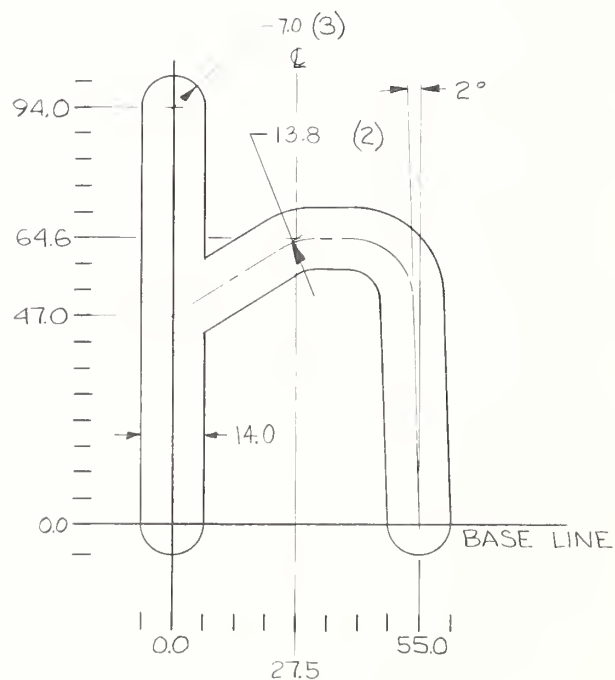


Figure II-75  
Small Letter H

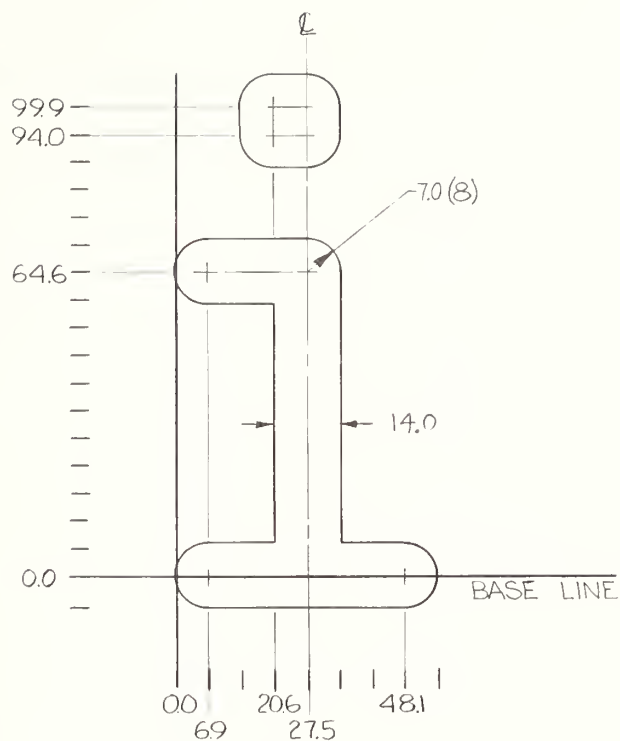


Figure II-76  
Small Letter I

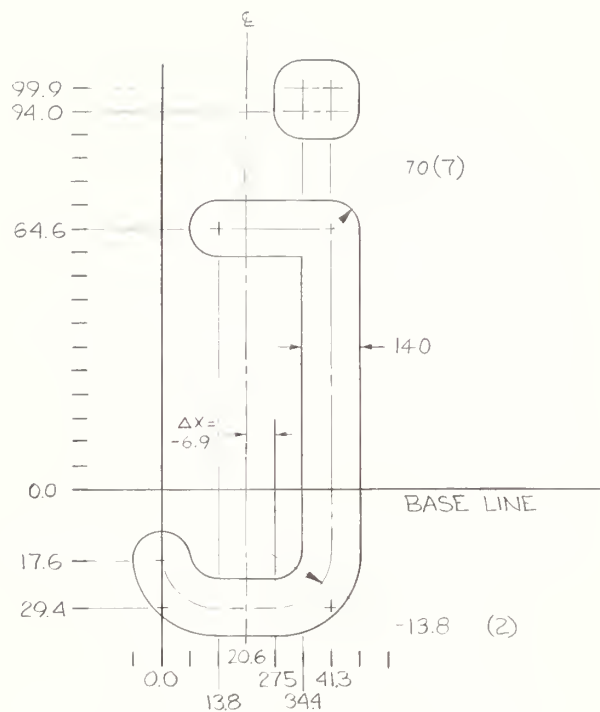


Figure II-77  
Small Letter J

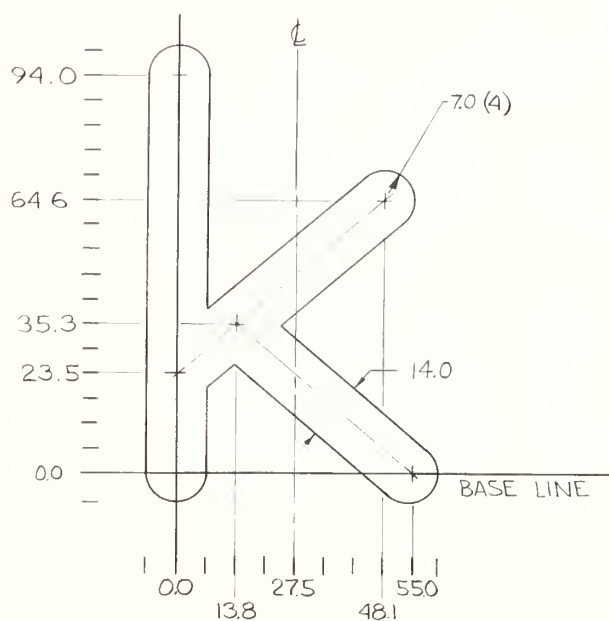


Figure II-78  
Small Letter K

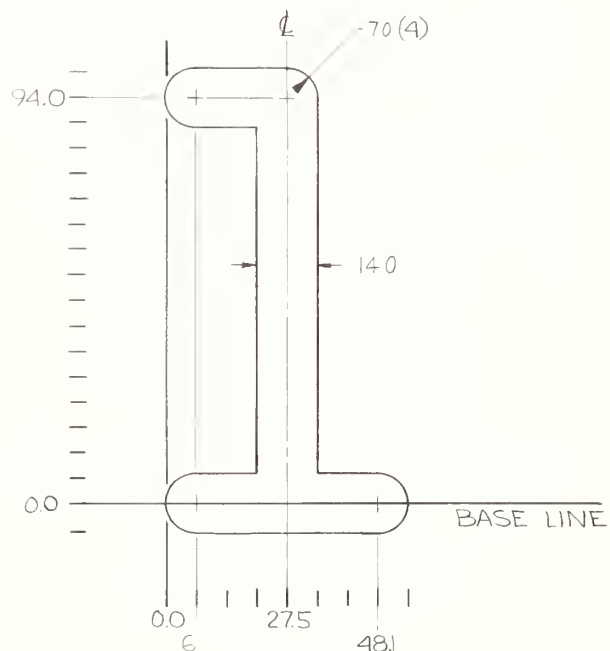


Figure II-79  
Small Letter L

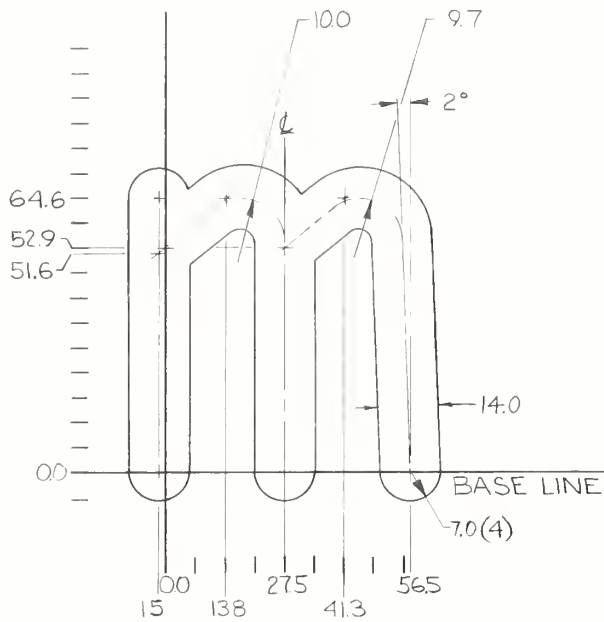


Figure II-80  
Small Letter M

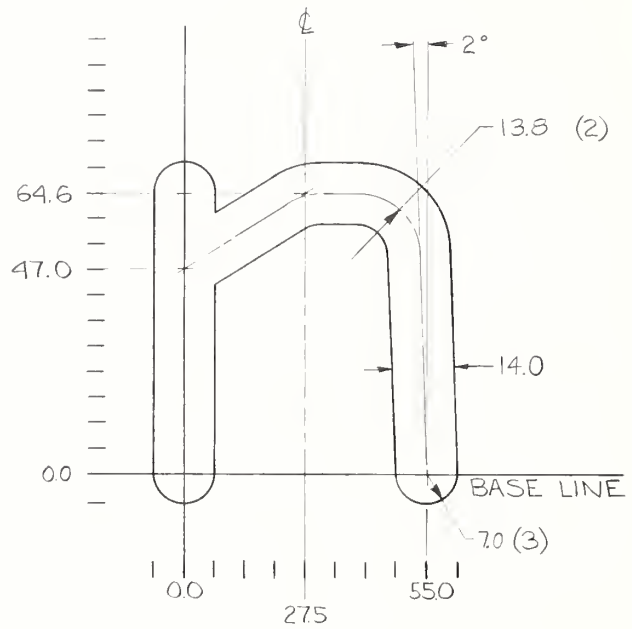


Figure II-81  
Small Letter N

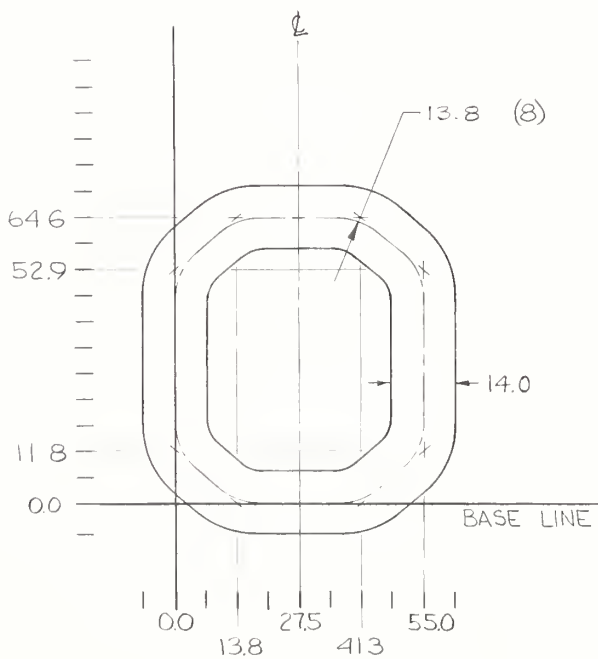


Figure II-82  
Small Letter O

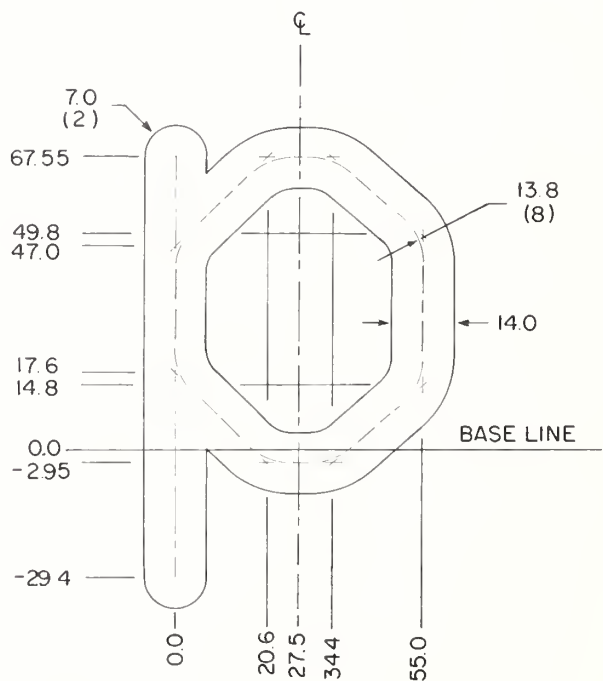


Figure II-83  
Small Letter P

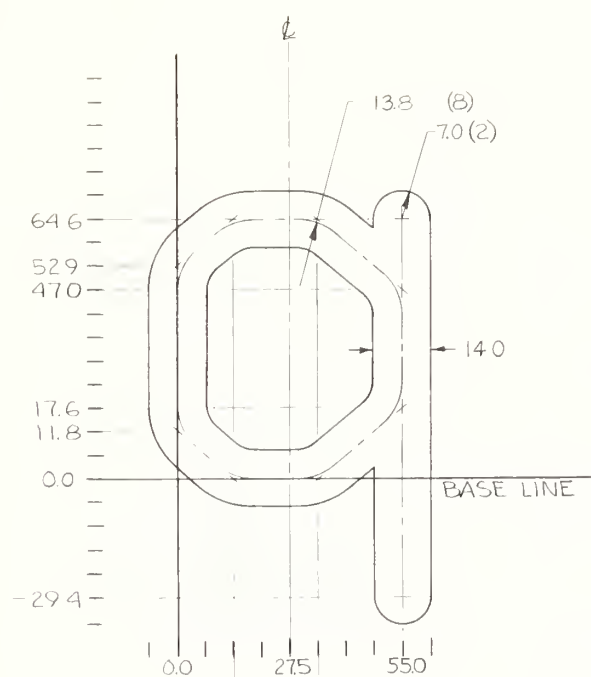


Figure II-84  
Small Letter Q

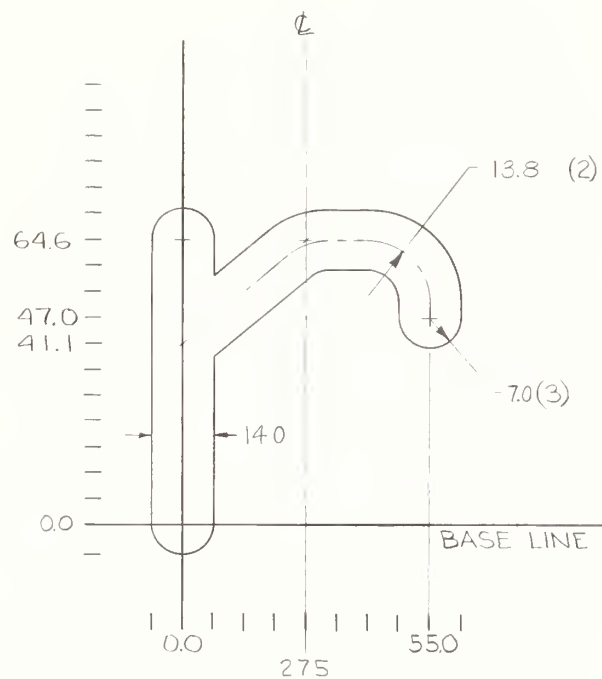


Figure II-85  
Small Letter R

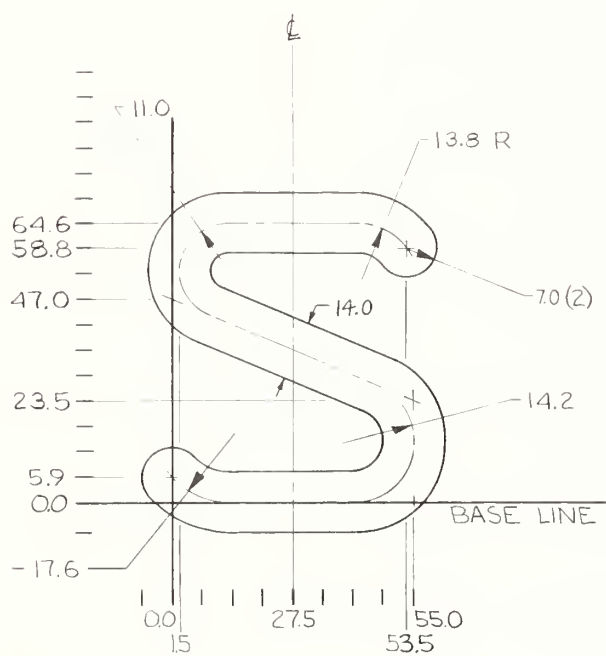


Figure II-86  
Small Letter S

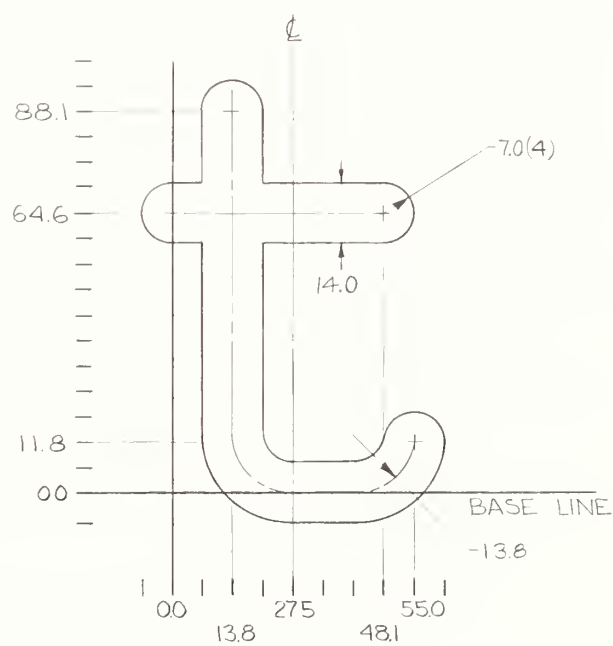


Figure II-87  
Small Letter T



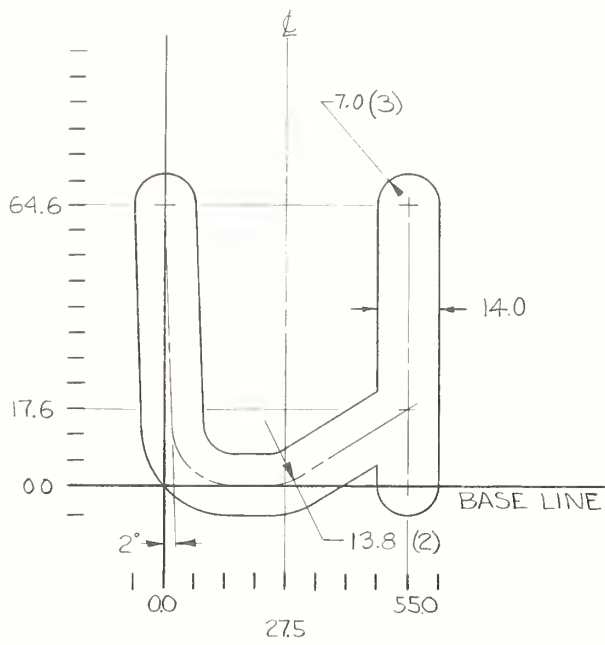


Figure II-88  
Small Letter U

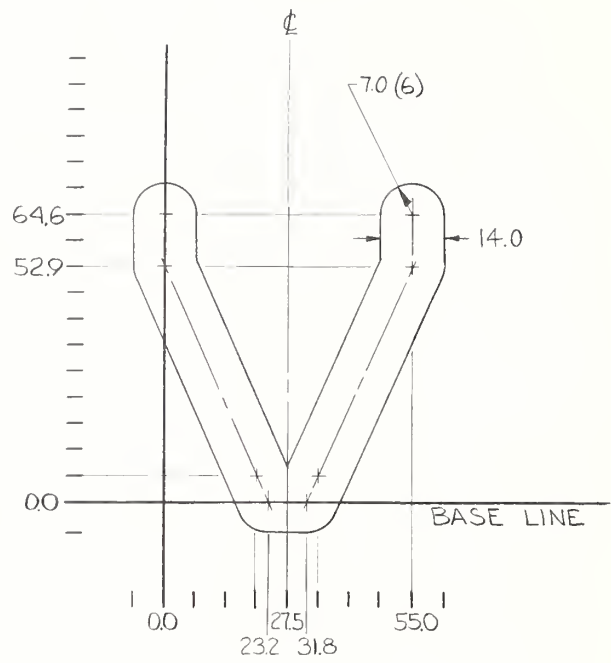


Figure II-89  
Small Letter V

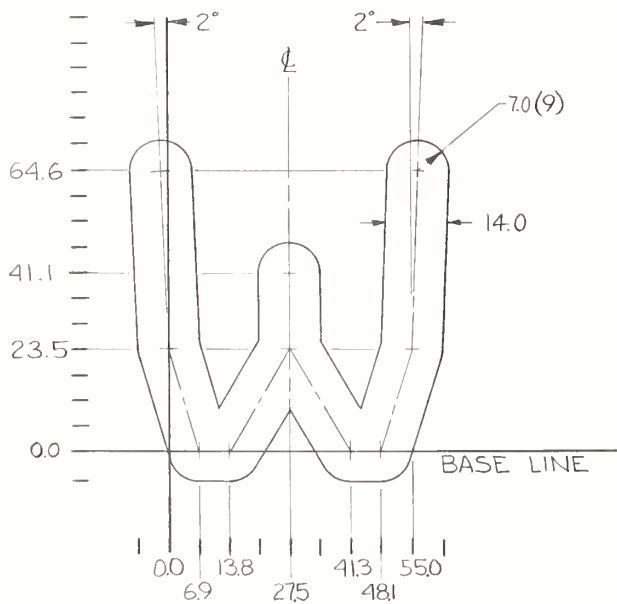


Figure II-90  
Small Letter W

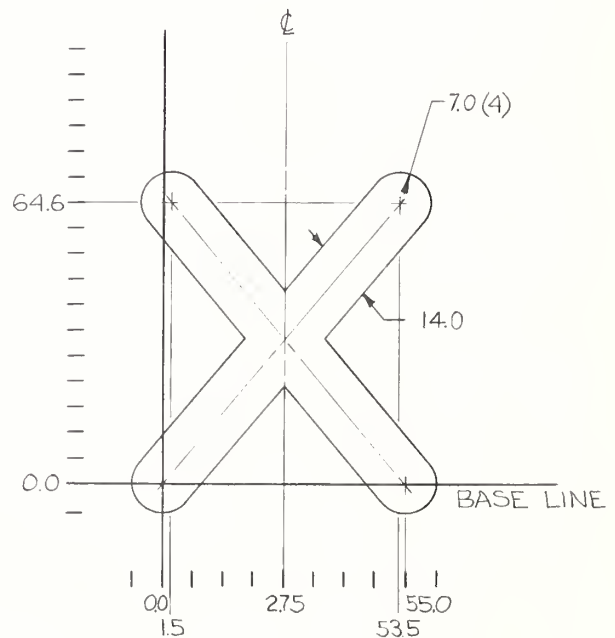


Figure II-91  
Small Letter X

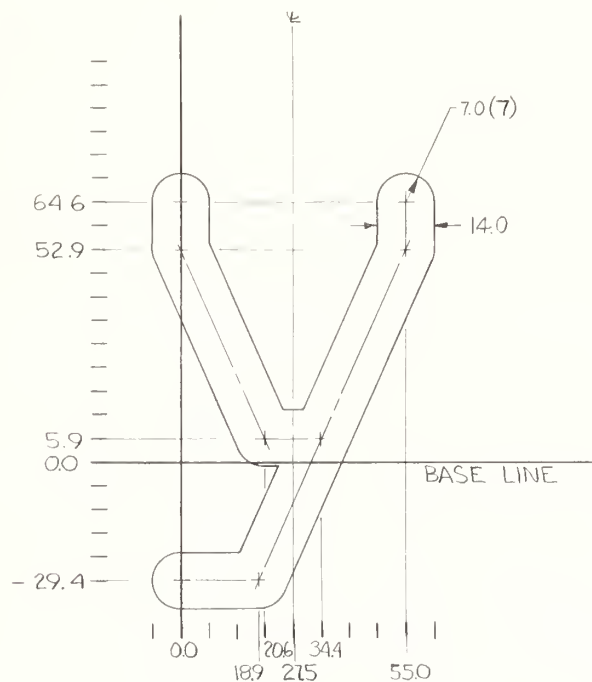


Figure II-92  
Small Letter Y

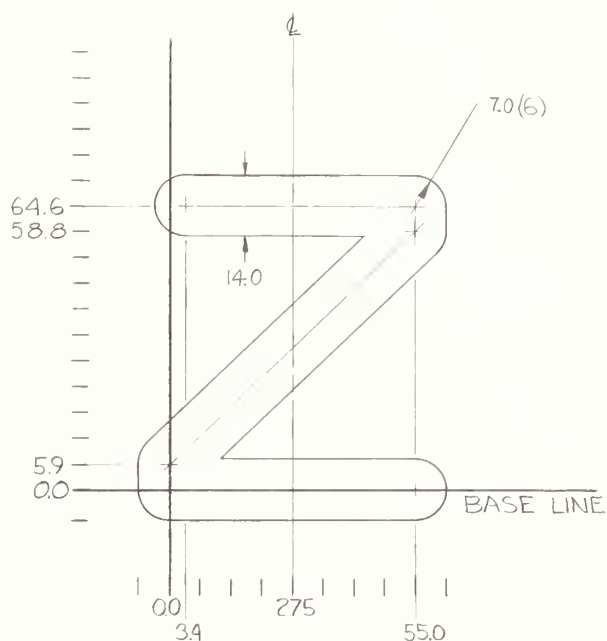
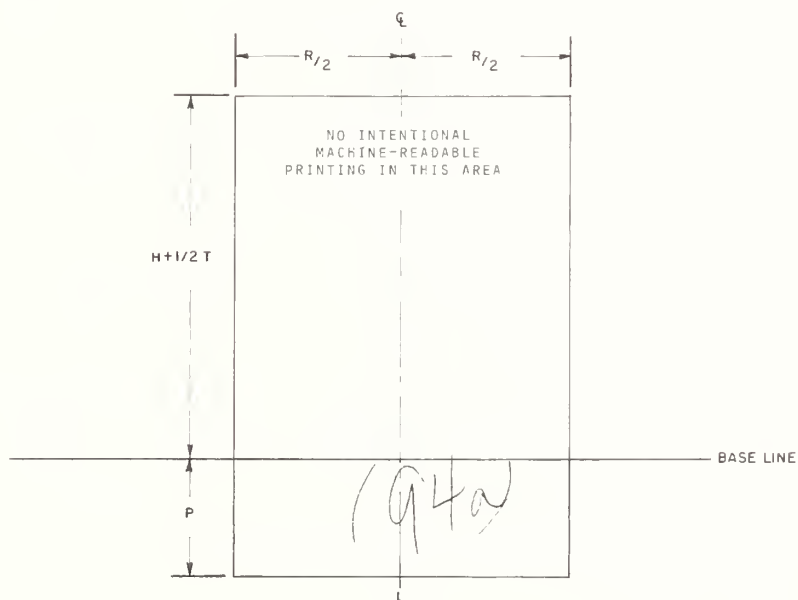


Figure II-93  
Small Letter Z



L IS THE NOMINAL CENTERLINE POSITION FOR A CHARACTER AS DEFINED BY THE PITCH OF THE PRINTING MECHANISM AND CHARACTERS ON THE LINE  
 $P = 5H/16 + T/2$  IF LOWER CASE IS USED, OR  $T/2$  IF LOWER CASE IS NOT USED  
 $R =$  NOMINAL CHARACTER SPACING (PITCH OF PRINTER),  $w + 2T$  MINIMUM

Figure II-94  
Character SPACE

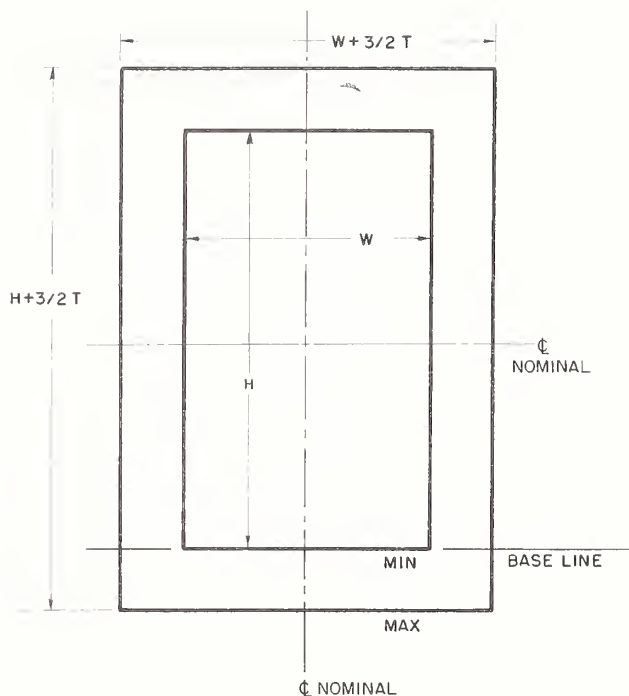


Figure II-95  
CHARACTER ERASE

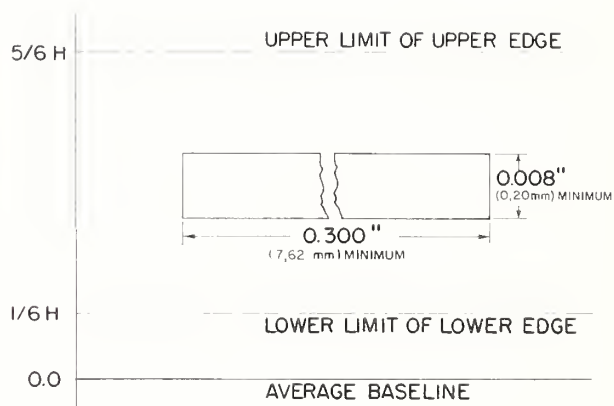


Figure II-96  
GROUP ERASE

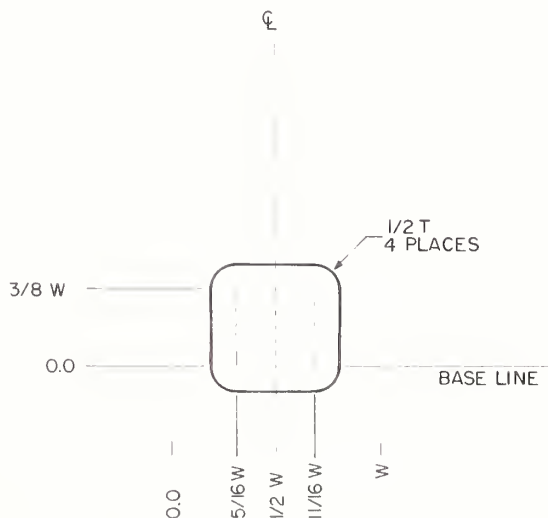


Figure II-97  
Alternate PERIOD

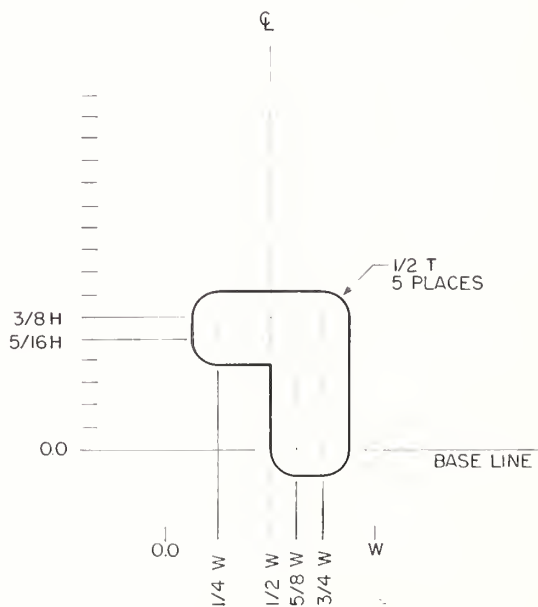


Figure II-98  
Alternate COMMA

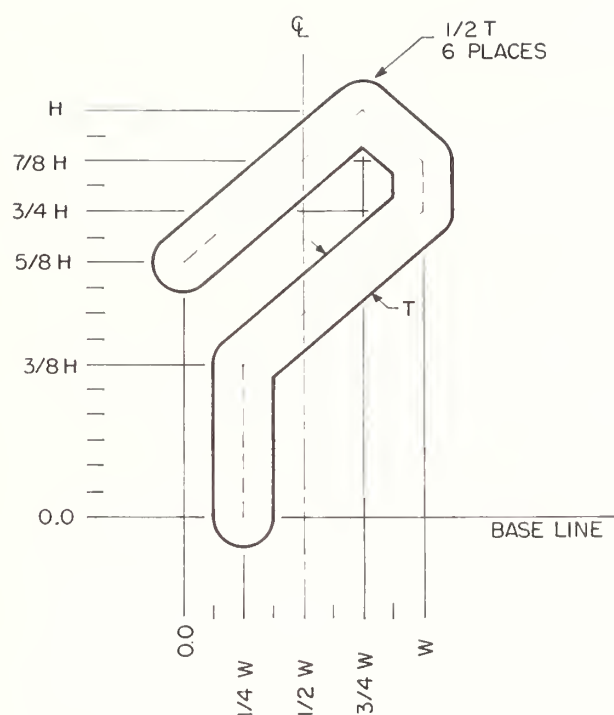


Figure II-99  
Alternate QUESTION MARK

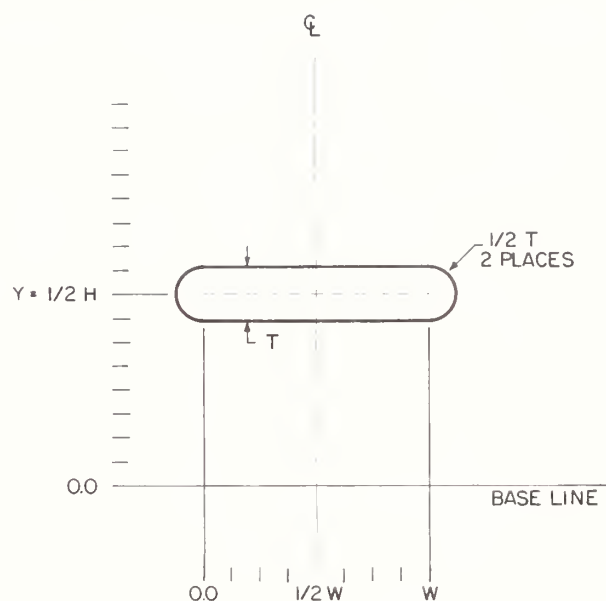


Figure II-100  
Alternate HYPHEN

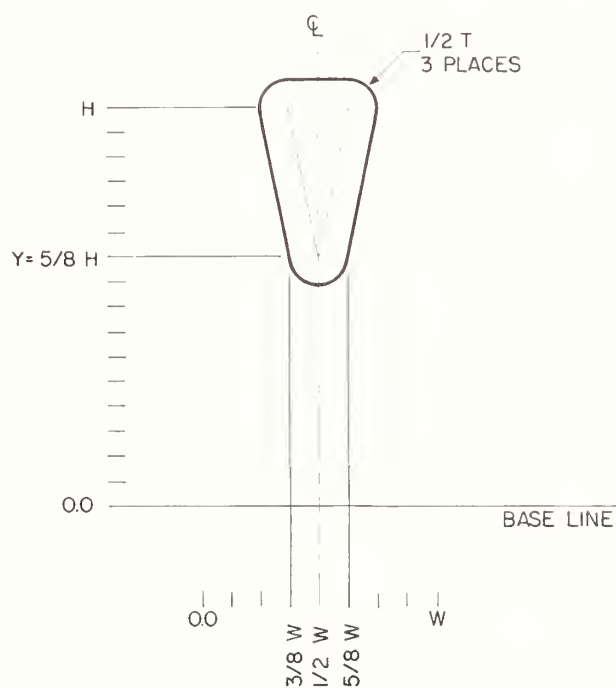


Figure II-101  
Alternate APOSTROPHE

## PART III—STYLE B

## 7. Standard Characters

7.1. **Character Sizes.** Standard character shapes are specified in three different sizes: I, III and IV. The applications for which these sizes were originally developed are similar to those given in Section 2.1.1, Font Size. Size III is for use with the numeric and journal tape subsets only. Relative size relationships with respect to centerline dimensions are given in Table III-1.

TABLE III-1—Relative Font Size Relationships—Style B

Size	Vertical	Horizontal
I.....	1.000	1.000
III.....	1.333	1.086
IV.....	1.500	1.500

Since the Style B sizes vary from character to character it is necessary to define each character size by scaling from precise master centerline drawings on stable material. These drawings are available as described in Subsection 7.9.2 and are reproduced in Section 9 out of scale for illustrative purposes only. Size I master centerline character drawings are superimposed on a coordinate grid of 2 mm resolution at a scale of 100:1. This represents a grid resolution of 0.000787 inch (0.02 mm) at full size. Size IV dimensions are derived from the same set of master drawings by magnifying the centerline dimensions by the factor 1.500. Stroke edges must be calculated with the assistance of the strokewidth information of Table III-3. Separate master centerline drawings are available for the numeric subset when used in Size III.

The largest character in overall size is the numeral ZERO. Its approximate centerline height and width is given in Table III-2.

TABLE III-2—Nominal Centerline Size for Numeral ZERO—Style B

Size	Nominal centerline height		Nominal centerline width	
	<i>Inch</i>	<i>(mm)</i>	<i>Inch</i>	<i>(mm)</i>
I.....	0.094	(2.40)	0.055	(1.40)
III.....	0.126	(3.20)	0.060	(1.52)
IV.....	0.141	(3.60)	0.083	(2.10)

7.1.1. **Strokewidth.** The nominal strokewidth for each size is given in Table III-3 below.

TABLE III-3—Nominal Strokewidth—Style B

Size	Nominal strokewidth lower case, #, % and @		Nominal strokewidth all other characters	
	<i>Inch</i>	<i>(mm)</i>	<i>Inch</i>	<i>(mm)</i>
I.....	0.012	(0.30)	0.014	(0.35)
III.....	Not available		0.015	(0.38)
IV.....	0.017	(0.43)	0.020	(0.50)



**7.2. Character Set Repertoire.** The printing graphics and character SPACE as defined in this standard constitute the total repertoire for optical character recognition. In some applications it may be desirable to use special characters herein defined for the purpose of error suppression or nondata functions.

There are no restrictions as to the information content of any OCR characters except character ERASE, GROUP ERASE and SPACE. The meaning of any character used in any particular application must be established by the user. Users are cautioned to ensure that there is a common understanding of the character sets employed in applications involving the interchange of documents.

**7.3. Subsets.** Subsets are NOT defined herein. They are the subject of a separate FIPS PUB. The user is cautioned that the complete repertoire may not be necessary, and an expanded set may adversely affect system performance. It is recommended that an appropriate minimum set be selected for each application.

**7.4. Relationship to ASCH Code Table.** Characters are defined herein for the entire character set of the FIPS PUB (1).

For OCR usage all characters are used in a stand-alone manner. Specifically, the UNDERLINE (DISCONTINUOUS), GRAVE ACCENT, UPWARD ARROWHEAD (CIRCUMFLEX), and OVERLINE stand as individual characters and are not combined with other characters to form composites.

The correspondence of SPACE, CHARACTER ERASE, GROUP ERASE and Long Vertical Mark are handled as follows:

SPACE is a normally nonprinting graphic character and corresponds exactly with ASCII character SPACE of Code Table Position 2/0.

CHARACTER ERASE and GROUP ERASE are format effectors in that the action of the reading machine is to ignore a character with CHARACTER ERASE superimposed upon it and to eliminate the line space otherwise occupied. GROUP ERASE elicits a similar action by the reading machine except that a group of characters are ignored. This action does not normally produce an output code. If the user must produce output coding for the CHARACTER ERASE or GROUP ERASE, they shall be transmitted as the ASCII character DELETE in Code Table Position 7/15.

Long Vertical Mark is a graphic character most generally associated with the function of field mark. It is usually used to delimit fields or data elements on OCR forms, particularly in applications in which the data is entered with keyboard driven devices. LVM can be associated for data transmission purposes with the ASCII character VERTICAL LINE of Code Table Position 7/12.

**7.5. Character SPACE.** The character SPACE is a blank area in a print line having a width equal to the width of the character pitch. The actual horizontal extent of the blank area between two horizontally adjacent characters depends on the number of SPACES included and on the width of the bounding characters. The accuracy with which the number of SPACE characters in a row can be determined depends upon the OCR scanner used, the print location tolerances and other factors. The width of multiple spaces and the response of a character reader to multiple spaces is, therefore, not covered in this standard.

**7.6. Character LONG VERTICAL MARK (LVM).** This character is normally used as a field separator and is usually distinguished from other characters by its unusually large vertical extent.

Minimum size for this character is given below. The use of this character is application dependent and the user is advised to consult his printer and OCR manufacturers.

TABLE III-4—Height of LONG VERTICAL MARK

LVM Size	Minimum Height	
	Inch	(mm)
I .....	0.146	(3.7)
III .....	0.196	(5.0)
IV .....	0.220	(5.6)

**7.7. CHARACTER ERASE.** The CHARACTER ERASE symbol has the special property that its presence is detectable when standing alone or when it is superimposed on any other printed character. It is intended to delete both the character that it covers and the line space that the character would otherwise occupy.

**7.8. GROUP ERASE.** GROUP ERASE is designed so that a long string of characters can be erased without striking a CHARACTER ERASE for each character to be deleted. It is defined as a continuous line between X and Y above the nominal base line, at least 0.300 inch long having a minimum thickness of 0.008 in (0.20 mm).

**7.9. Character Shapes and Dimensions.** The character shapes are defined by precise master centerline drawings on stable material. The procedure for obtaining accurate stable copies is given in Subsection 7.9.2. Paper reproductions of the drawings are also available for use when precision of scale is unimportant.

The drawings show the centerlines of the character strokes. The full character comprises the area covered by a circle of diameter equal to the strokewidth which is placed with its center on the character centerline and is made to traverse the entire extent of the center line. In the vicinity of stroke endings or intersections there may be exceptions to the general rule. All stroke edges in the vicinity of stroke endings and intersections are shown on the master drawings (see Subsection 7.9.1, below).

The Size I drawings are superimposed on an accurate rectangular grid which permits digitalization of the character shapes if desired. The resolution of the grid at full size is 0.000787 inch (0.0200 mm).

**7.9.1. Special Considerations for Style B.** The drawings show external square corners on characters such as B, D, E, F, G and so forth. It is important for reliable OCR performance, especially on B and D, that these corners not be rounded. It is advised that special attention be given to this in the design of type.

**7.9.2. Procedure for Obtaining Duplicate Stable Drawings of Style B Characters.** Duplicates of the centerline drawings on a stable base at exact 100:1 scale on a 280 mm × 380 mm grid can be obtained upon request. Paper reproductions are also available. Their quality is such

that they should not be further reproduced. Indicate if Size I or Size III style B font is desired. Size IV can be derived from Size I (see Section 7.1).

Address: Computer Systems Engineering Division  
Institute for Computer Sciences and Technology  
National Bureau of Standards  
Washington, D.C. 20234

**7.9.3. Letterpress Version of OCR-B.** The Style B characters are defined in this standard to have essentially constant strokewidths. This design allows for a maximum of deterioration of the quality of the printed image while still maintaining OCR separability. The ECMA European Standard includes a second version of the font which may be used with very high quality printing processes (letterpress, for example). This version is based on the identical centerline description but the strokewidths vary and stroke endings are specially designed. The objective is to improve the appearance of the font for printing processes which are extremely accurate.

This letterpress version is *not part* of this FIPS publication for Optical Character Recognition and the user is referred to document ECMA-11, 2d Edition, October 1971 for illustrations and further detail. It may be obtained from: ECMA, 114 Rue du Rhone, 1204 Geneva, Switzerland.

## 8. Character Positioning

**8.1. Format Rules.** Character positioning specifications are needed to insure that each OCR character is seen by the reading device without interference from other OCR characters or from non-OCR matter. The rules which define the form reference edges, clear area, printing area, margin and data fields are the same as those of Style A and may be found in Subsections 3.2 thru 3.6. Character spacing and line separation standards are given below in Subsections 8.2 and 8.3. These sections contain basic specifications relating to the position of characters on a form to accommodate the general requirements of OCR devices. It does not contain all of the rules which may be necessary for a particular application.

**8.2. Character Spacing.** Each standard drawing has indicated upon it indexing marks to indicate a horizontal base line (◄) and a print position centerline (▲).

A row of characters is properly aligned when all the base lines are collinear. Characters may be spaced horizontally either uniformly (constant pitch) or nonuniformly (proportionally spaced). For constant pitch printing the character centerlines are spaced at a distance of at least 0.0833 inch (2.14 mm) for Size I and 0.143 inch (3.63 mm) for Sizes III and IV. For proportionally spaced printing adjacent characters are separated by a horizontal blank area of at least one nominal strokewidth in extent. The spacing of the center lines depends not only on this value but on the tolerances with which the width of the characters is maintained and with which the relative positioning is held.

It is advisable to check with the OCR manufacturer when considering any centerline spacing less than 0.100 in (2.55 mm). A more economical reader may be obtained with 0.100 inch (2.55 mm) spacing.

8.3. **Line Separation.** The minimum distance from the lowest vertical extension of one line of characters to the highest extension of the next lower line is given below in Table III-5.

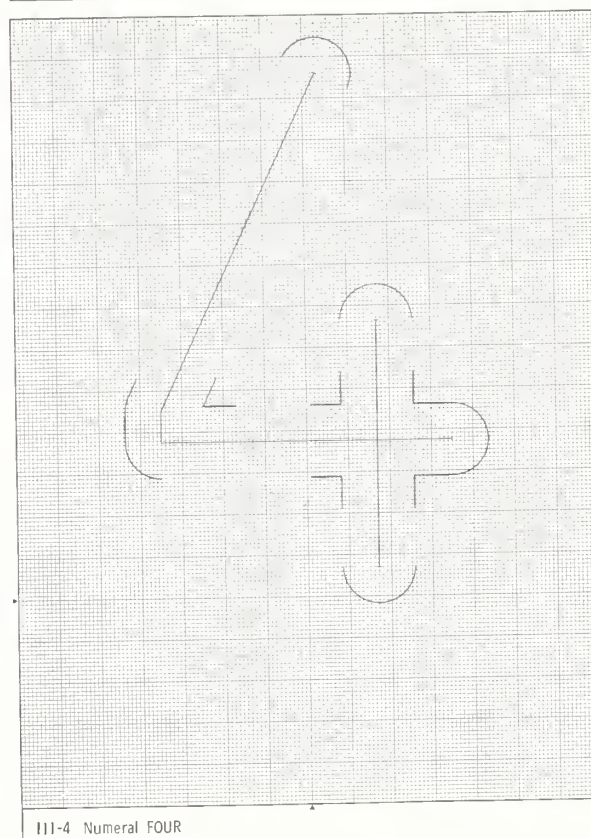
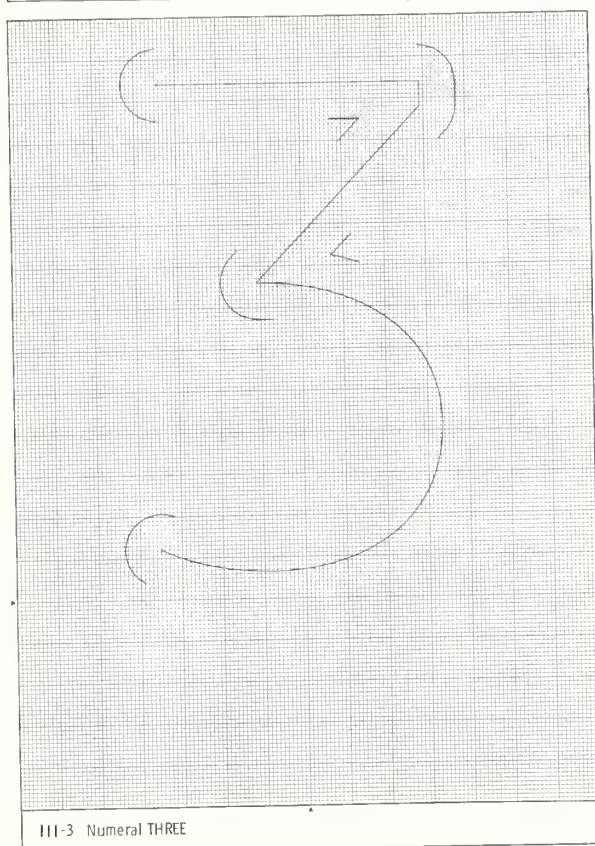
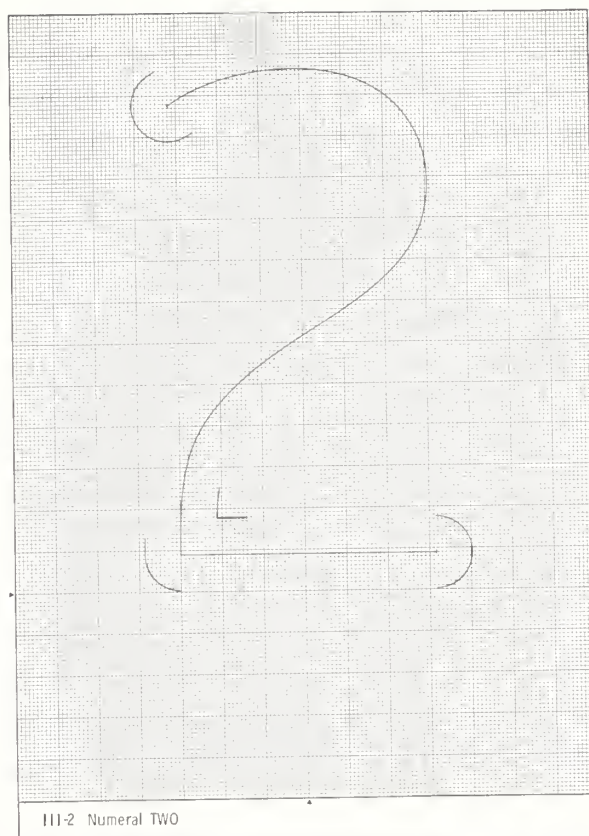
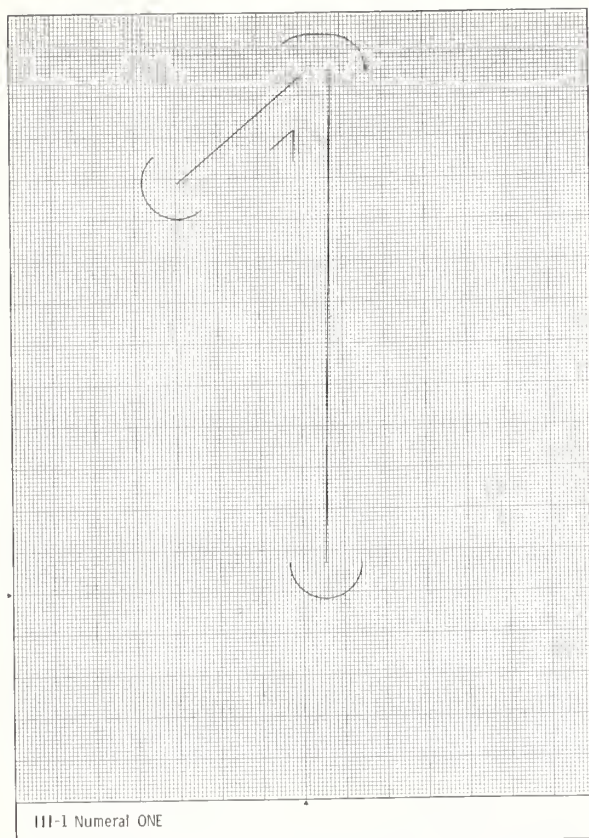
TABLE III-5—Minimum Line Separation—Style B

Size	Minimum line separation	
	<i>Inch</i>	<i>(mm)</i>
I.....	0.025	(0.64)
III.....	0.060	(1.52)
IV.....	0.080	(2.03)

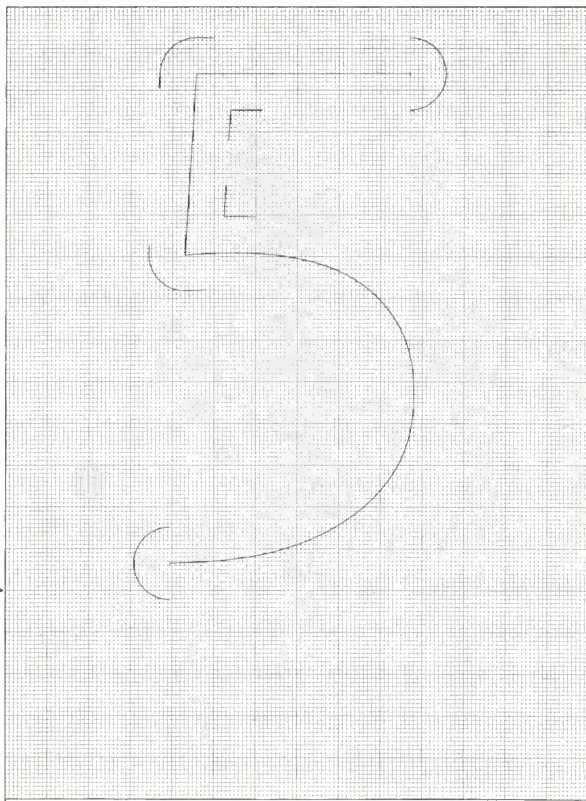
## 9. Individual Character Centerline Drawings

The following drawings are for illustrative purposes only and are not to scale. For stable master centerline drawings see Subsection 7.9.2.

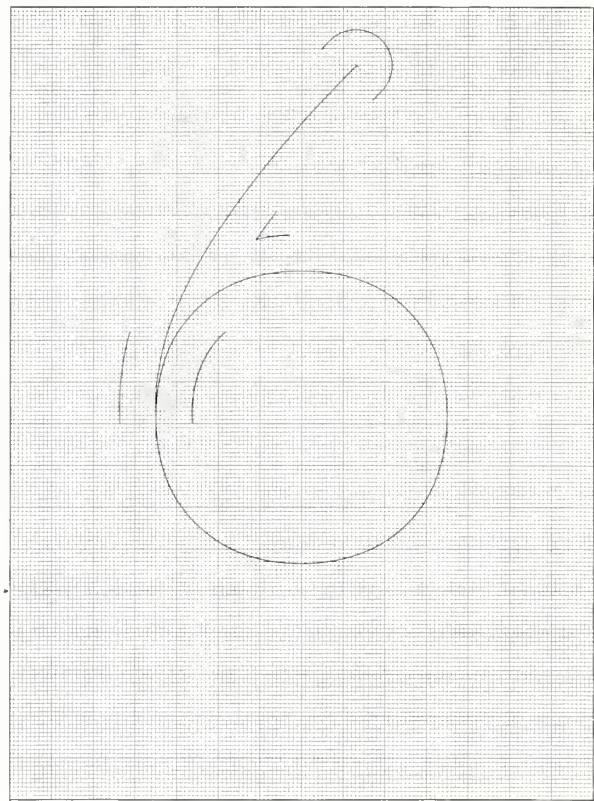




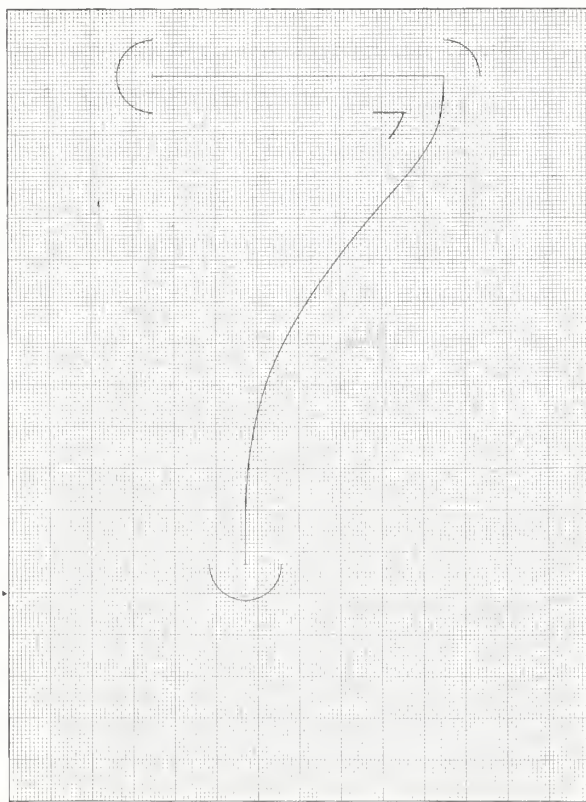




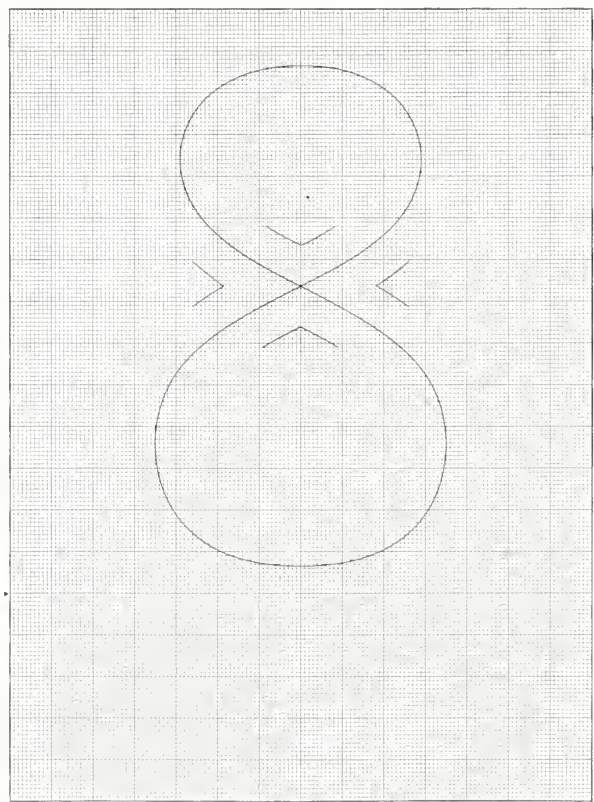
III - 5 Numeral FIVE



III - 6 Numeral SIX

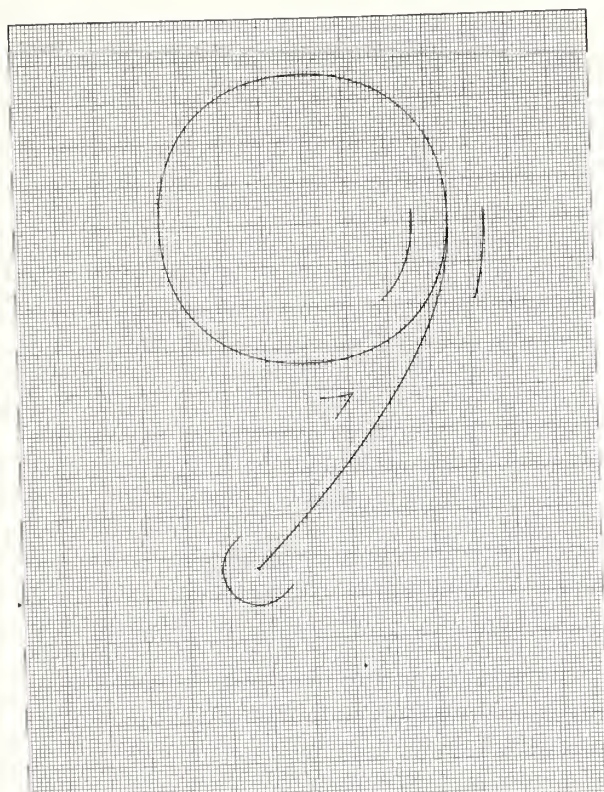


III - 7 Numeral SEVEN

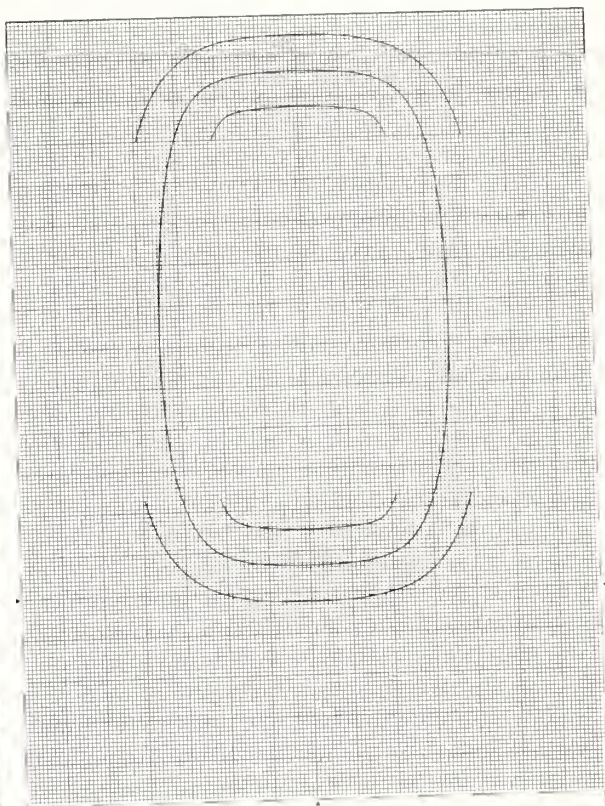


III - 8 Numeral EIGHT

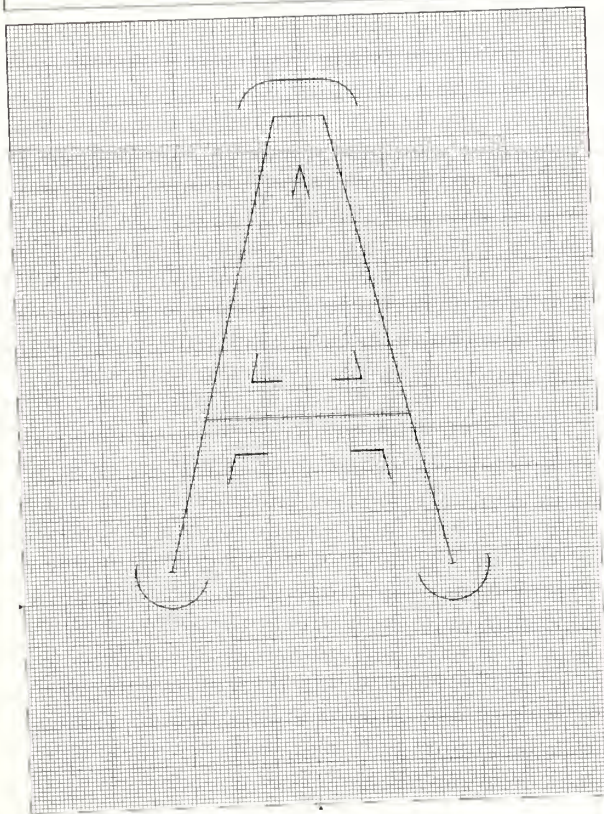




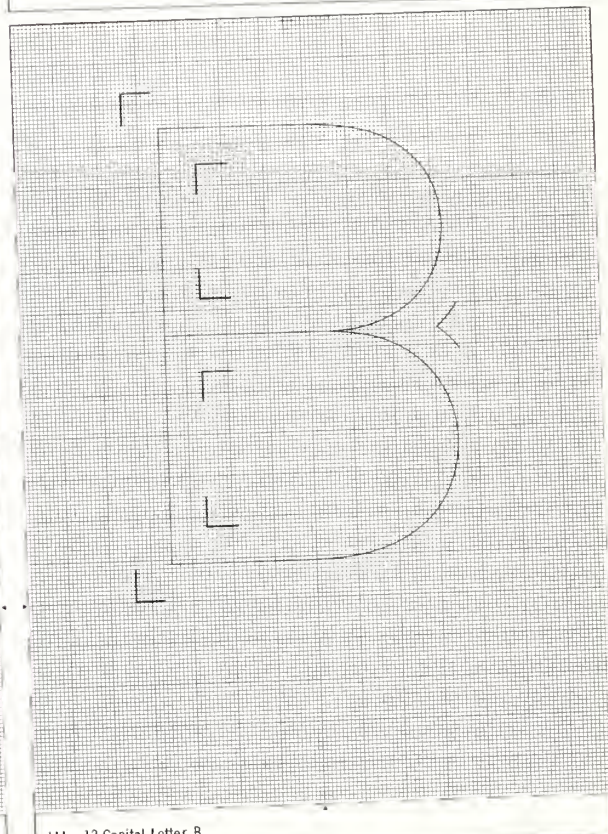
III - 9 Numeral NINE



III - 10 Numeral ZERO

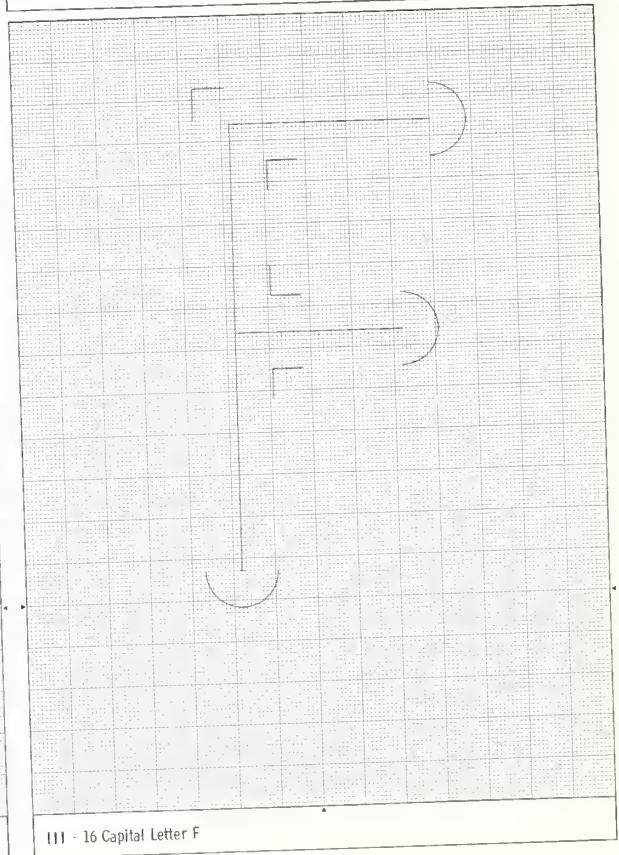
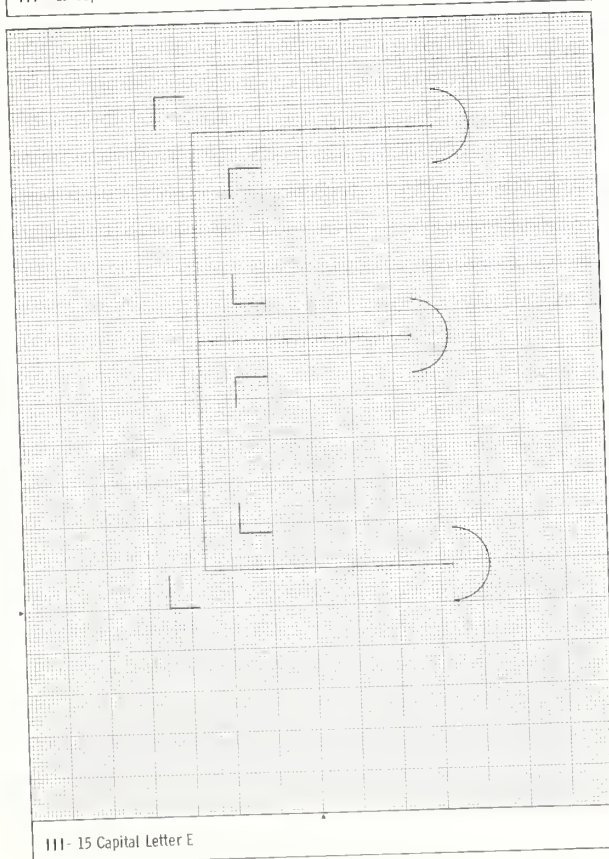
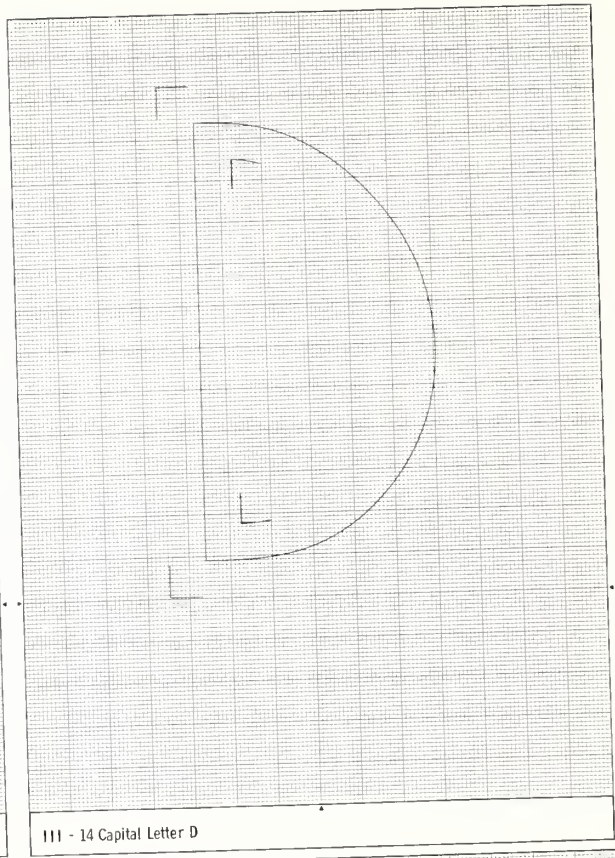
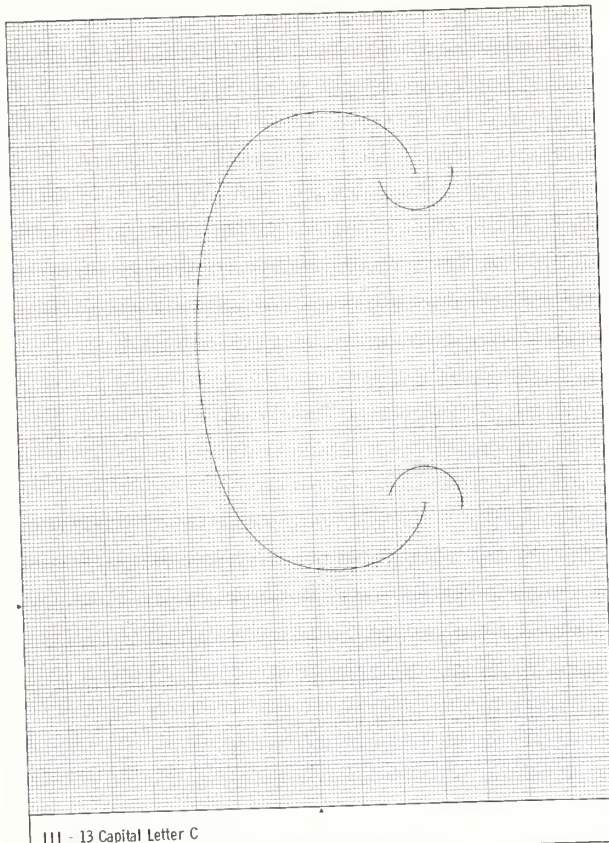


III - 11 Capital Letter A

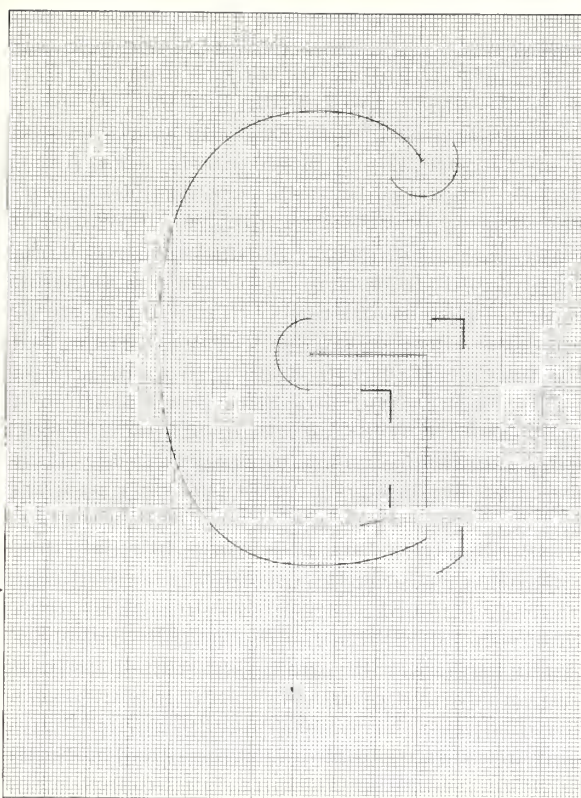


III - 12 Capital Letter B

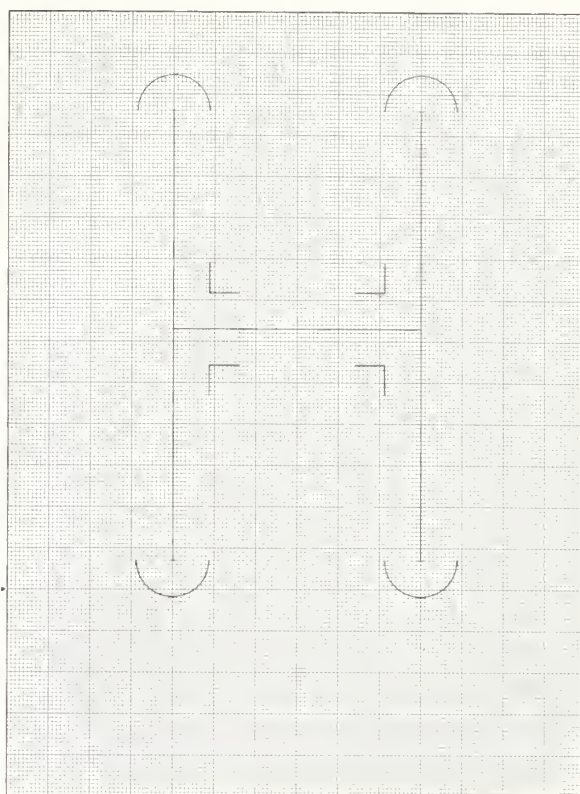




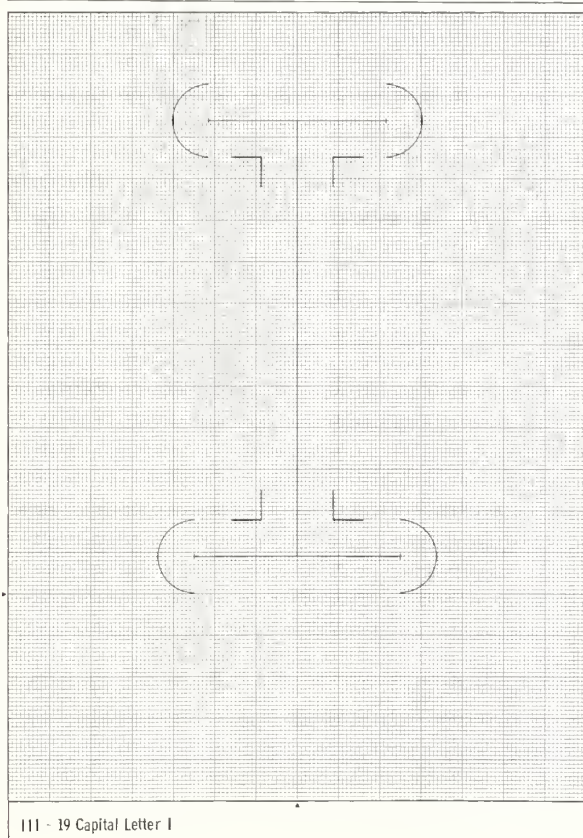




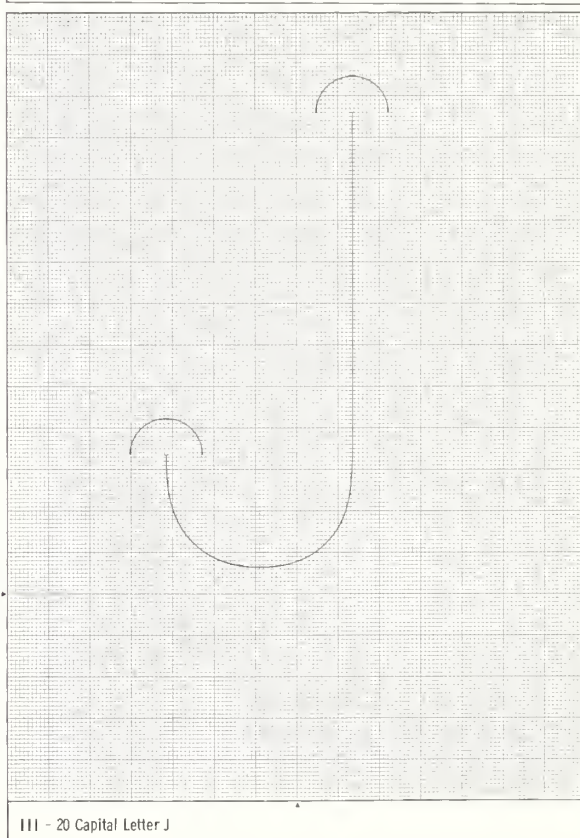
III - 17 Capital Letter G



III - 18 Capital Letter H

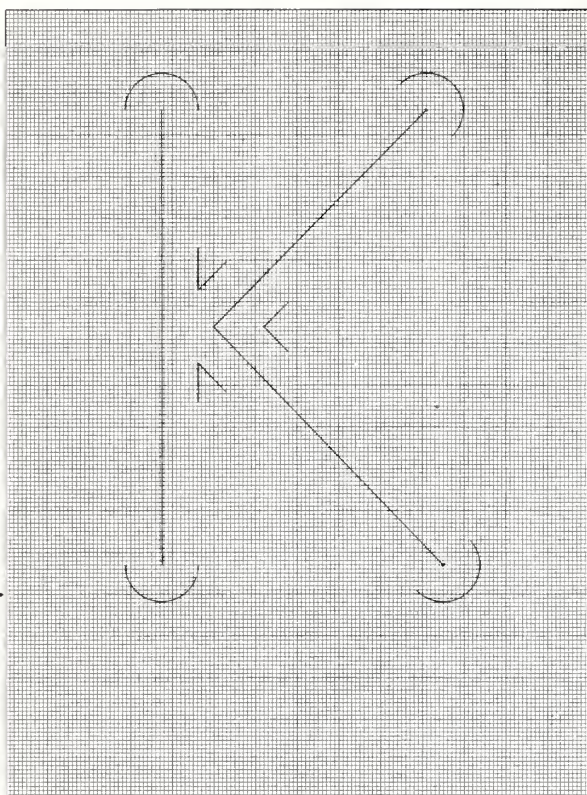


III - 19 Capital Letter I

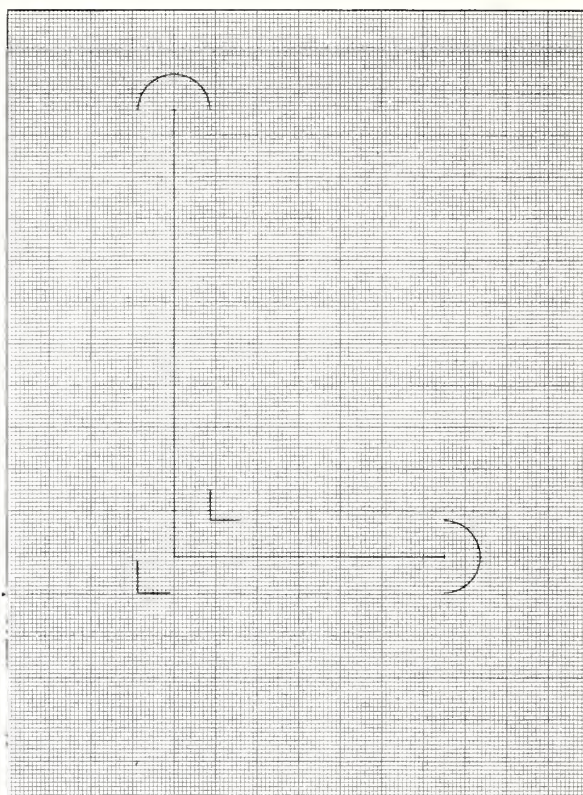


III - 20 Capital Letter J

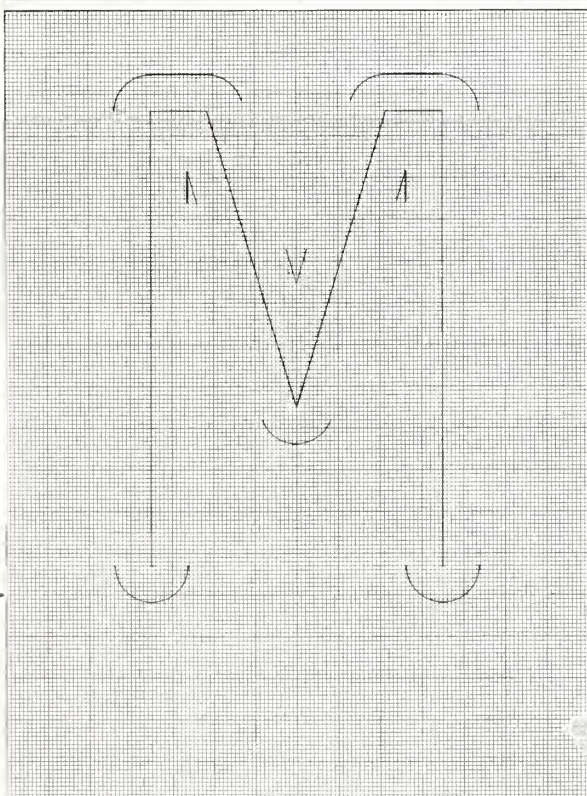




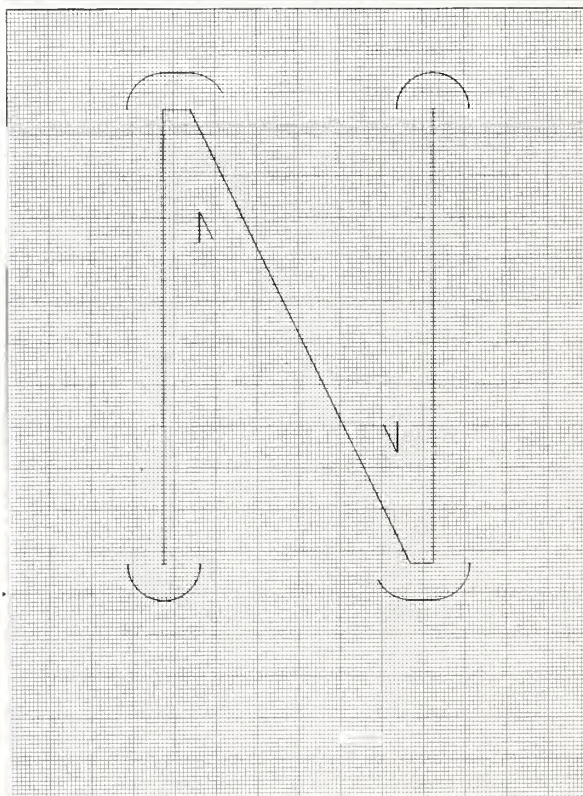
III - 21 Capital Letter K



III - 22 Capital Letter L

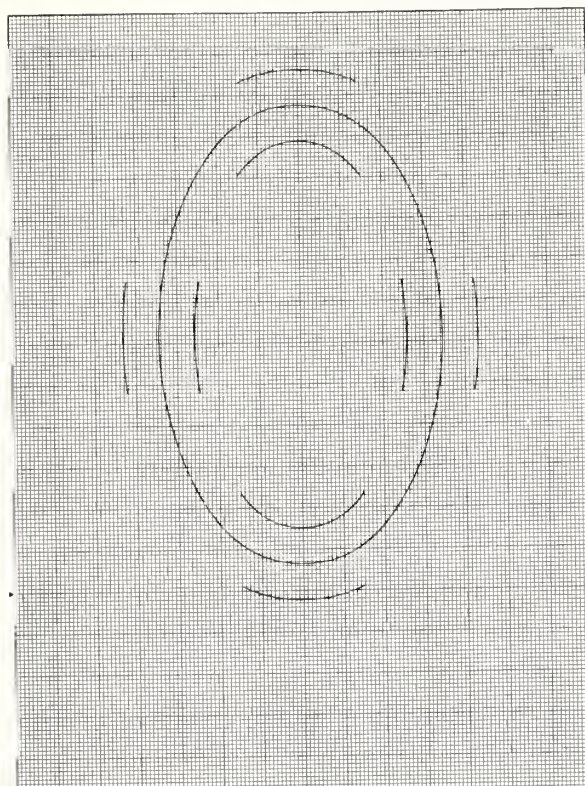


III - 23 Capital Letter M

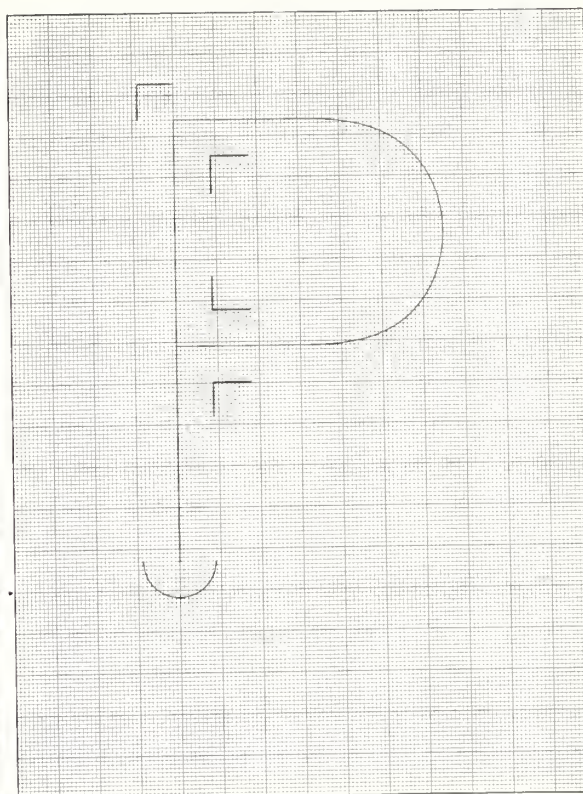


III - 24 Capital Letter N

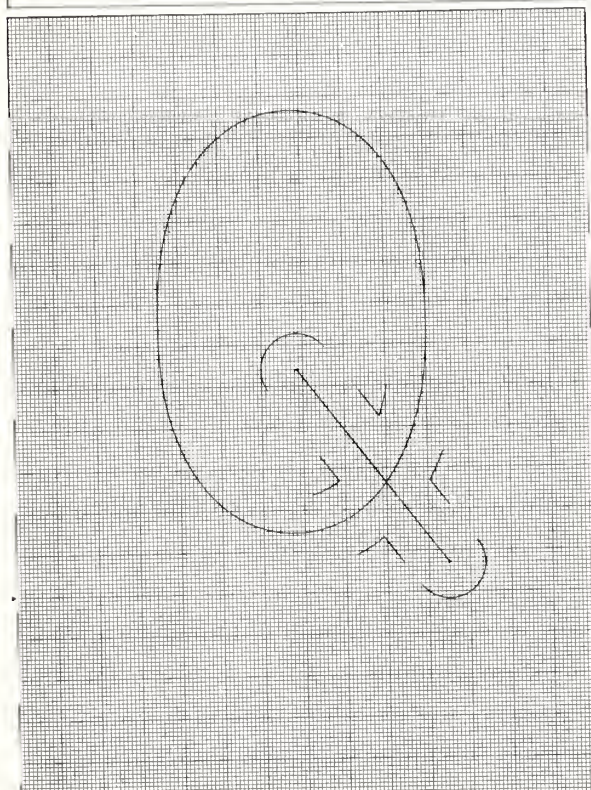




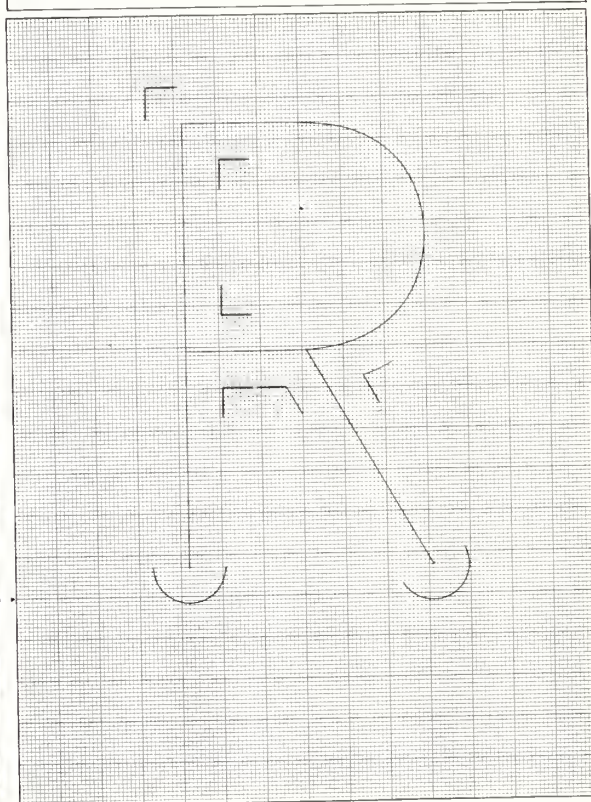
III - 25 Capital Letter O



III - 26 Capital Letter P

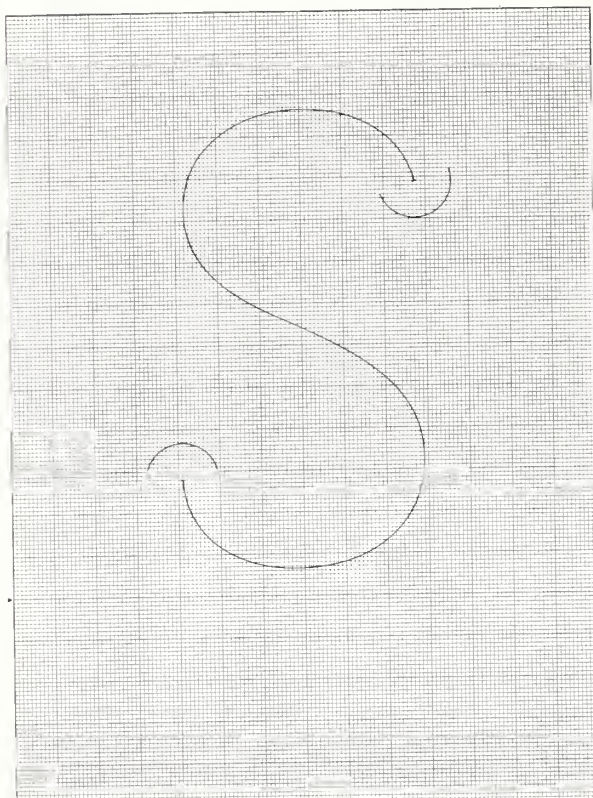


III - 27 Capital Letter Q

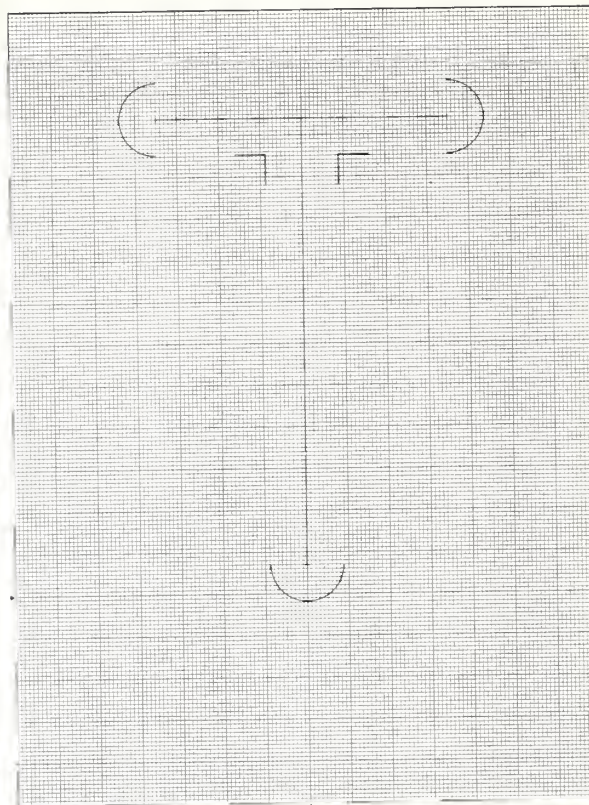


III - 28 Capital Letter R

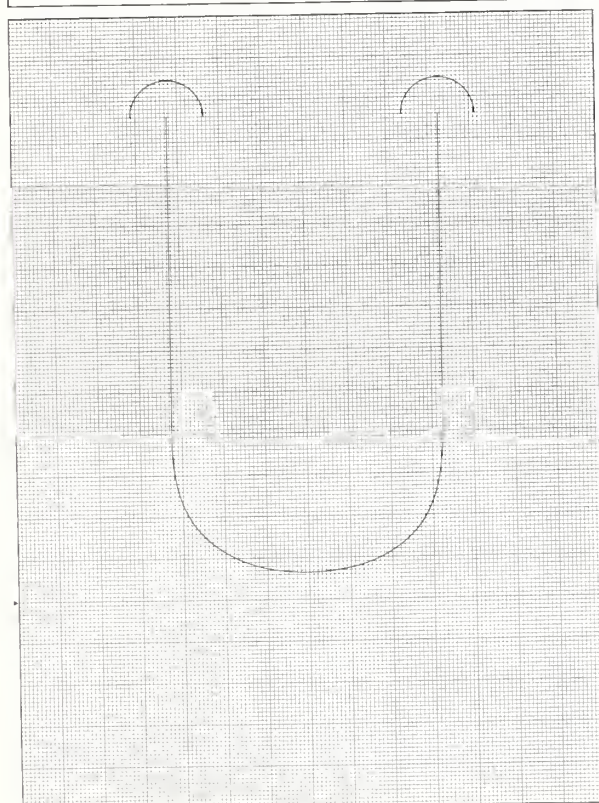




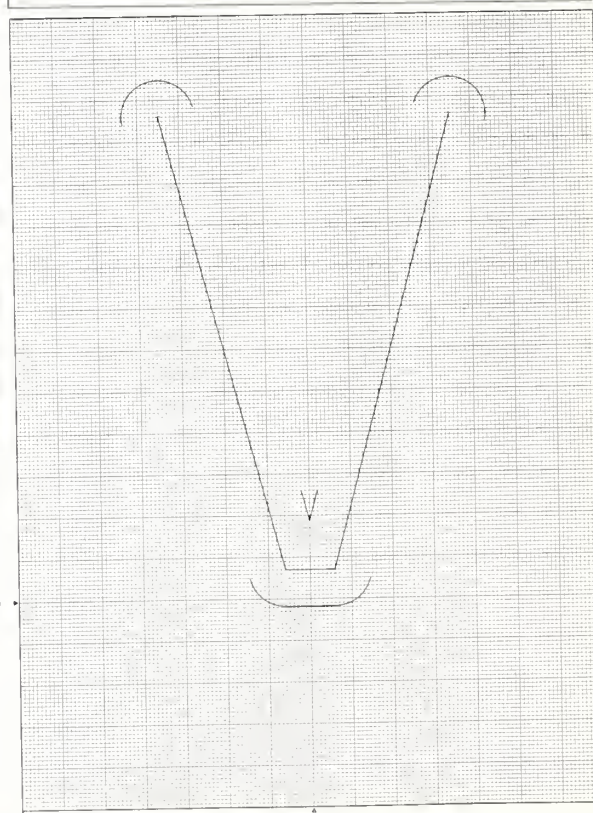
III - 29 Capital Letter S



III - 30 Capital Letter T

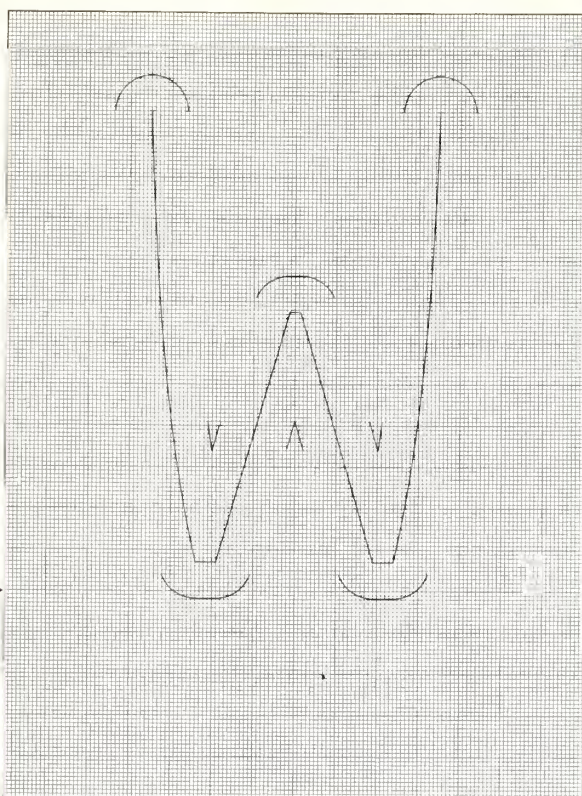


III - 31 Capital Letter U

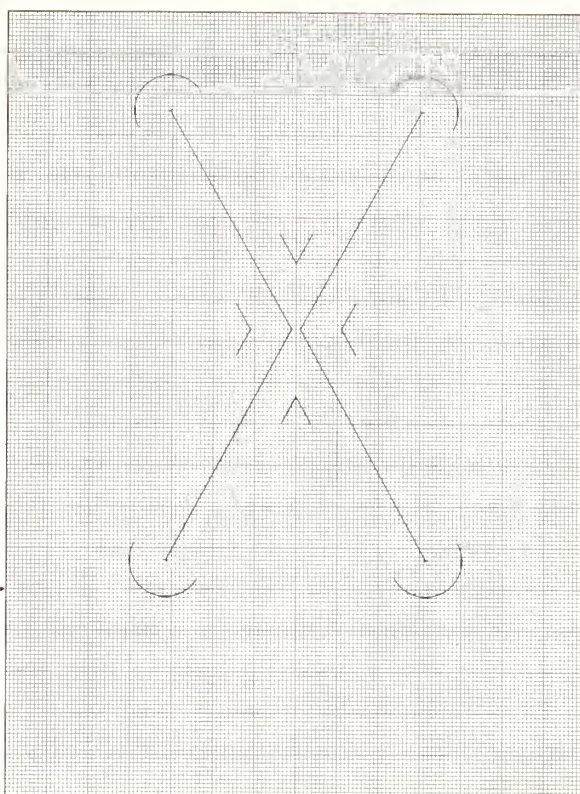


III - 32 Capital Letter V

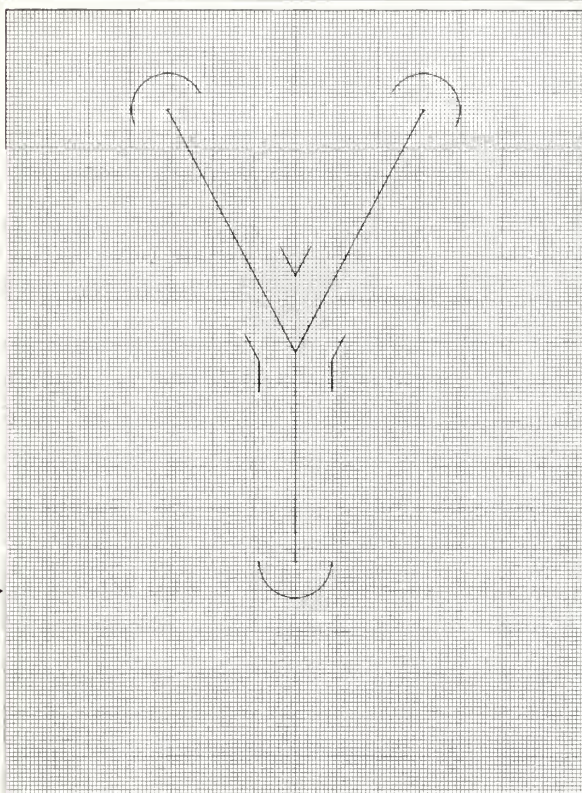




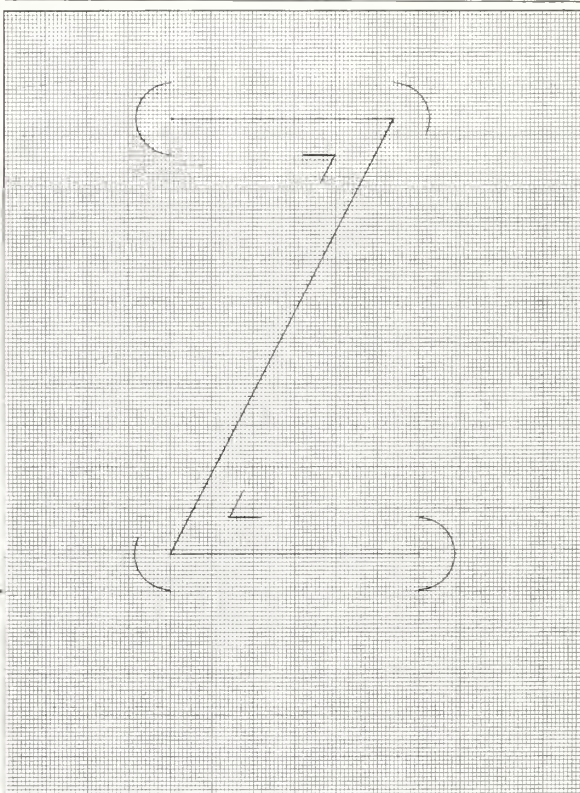
III - 33 Capital Letter W



III - 34 Capital Letter X

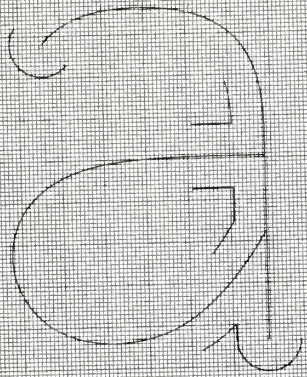


III - 35 Capital Letter Y

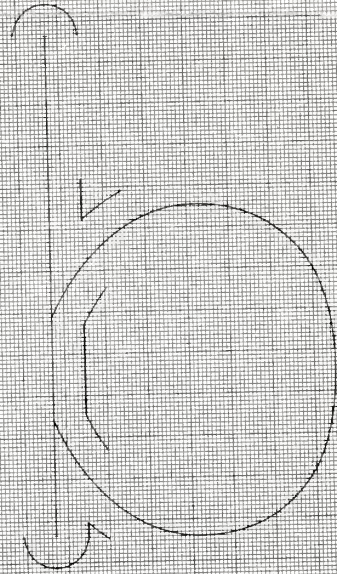


III - 36 Capital Letter Z

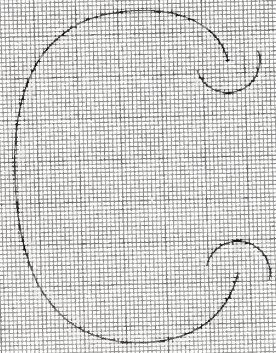




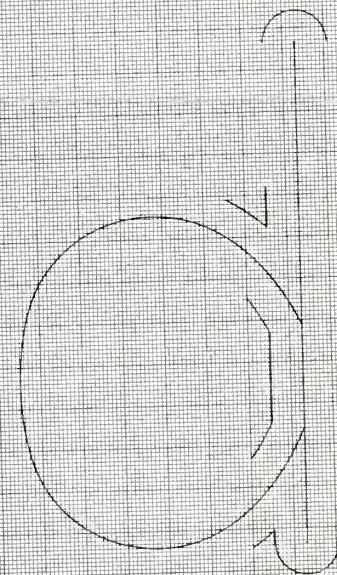
III - 37 Small letter a



III - 38 Small letter b

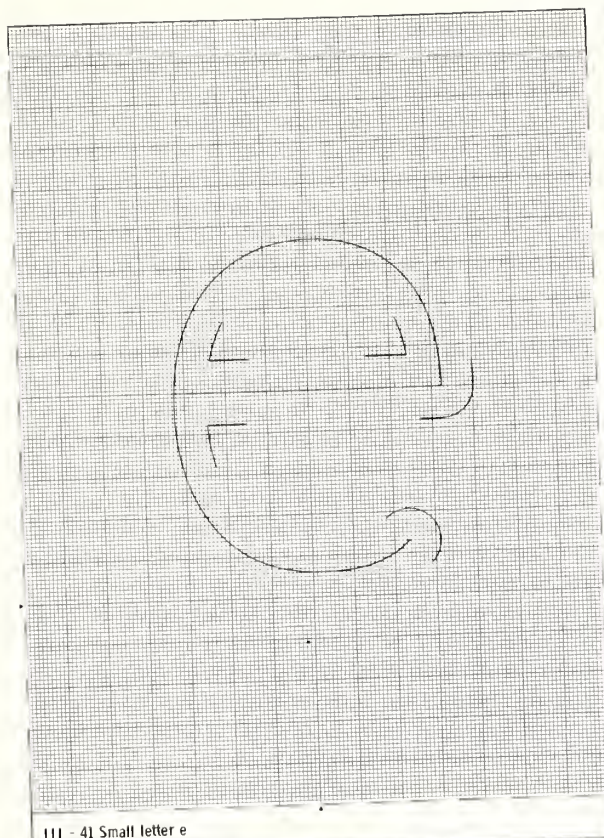


III - 39 Small letter c

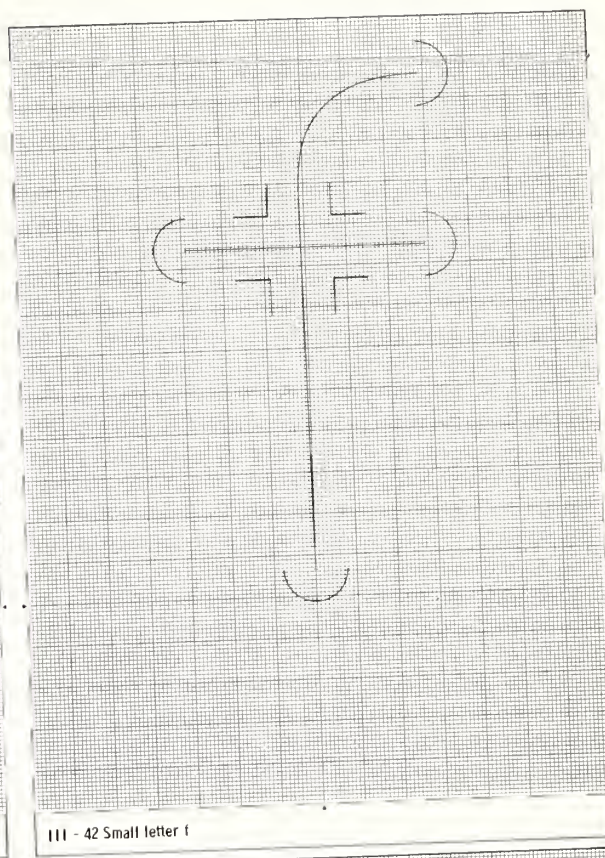


III - 40 Small letter d

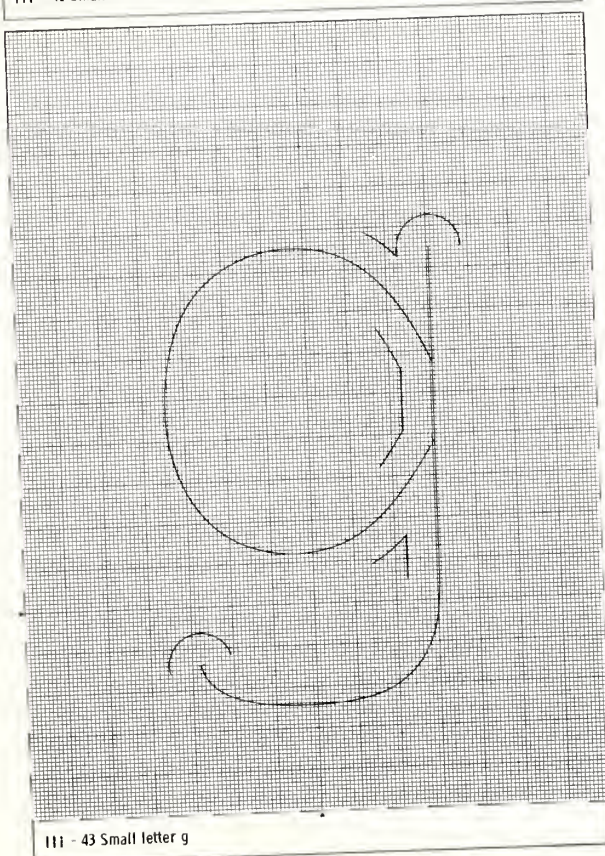




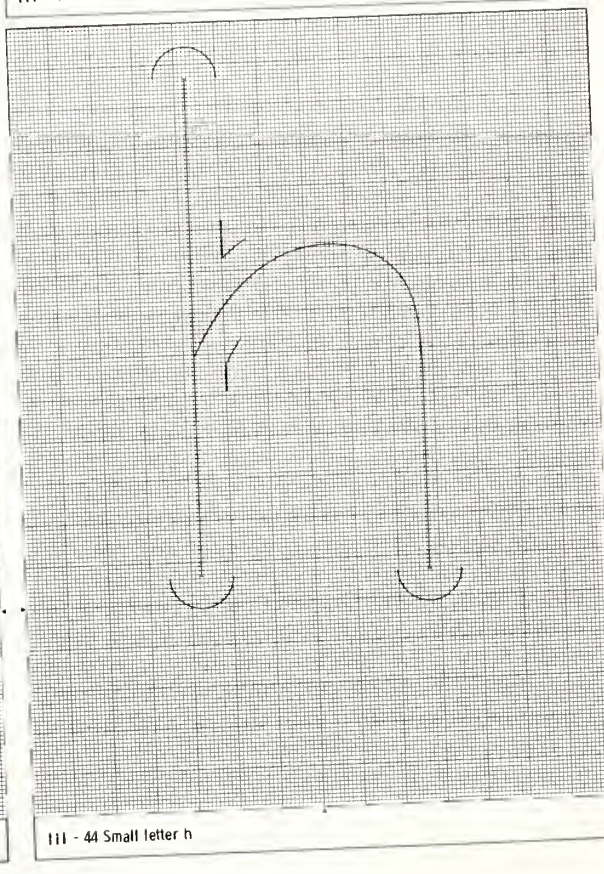
111 - 41 Small letter e



111 - 42 Small letter f

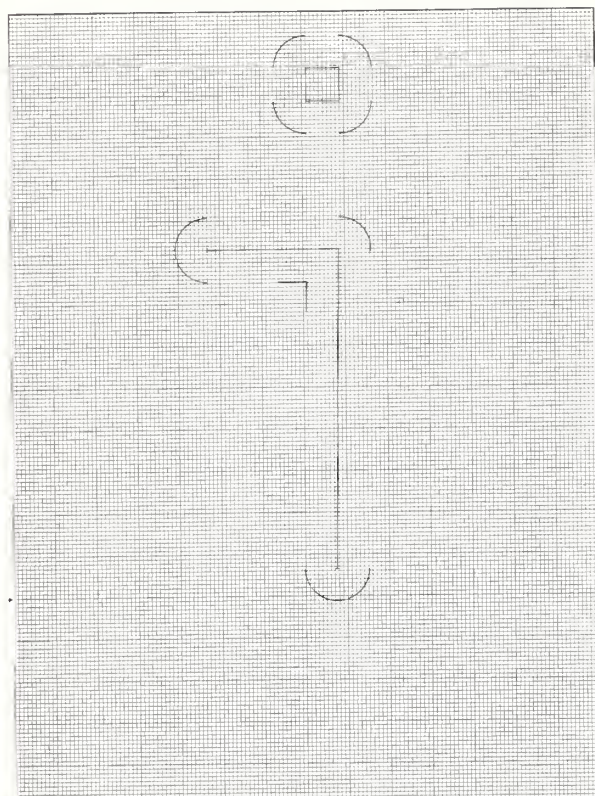


111 - 43 Small letter g

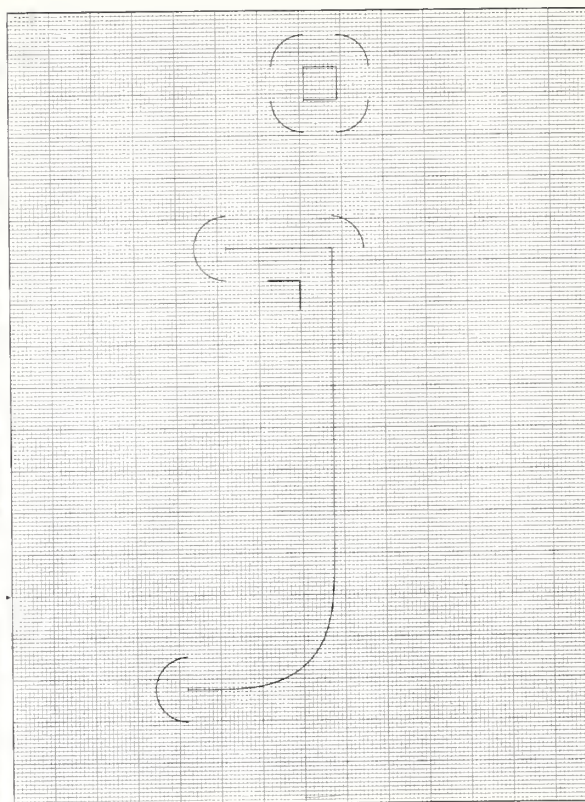


111 - 44 Small letter h

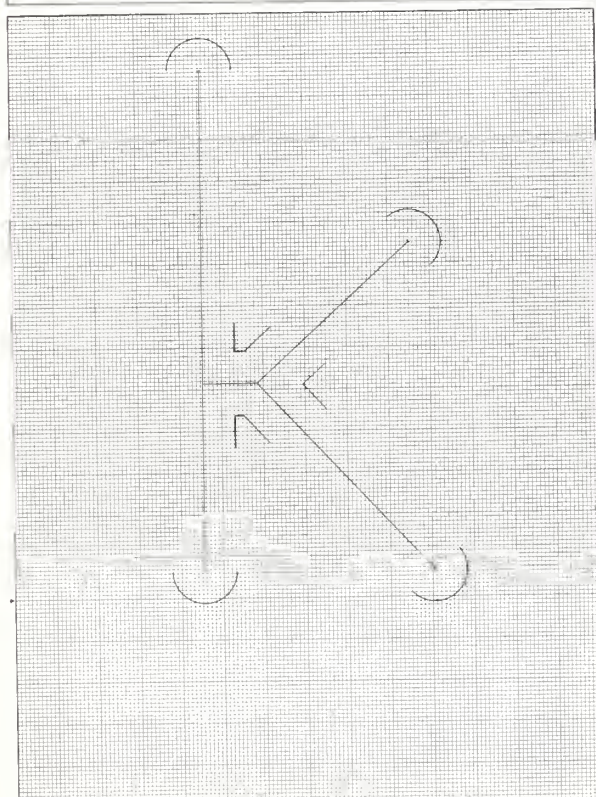




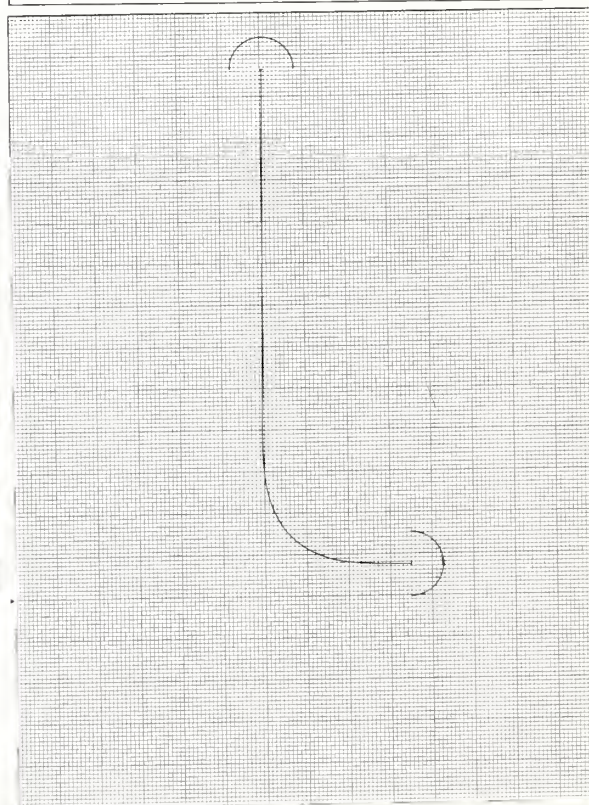
III - 45 Small letter i



III - 46 Small letter j

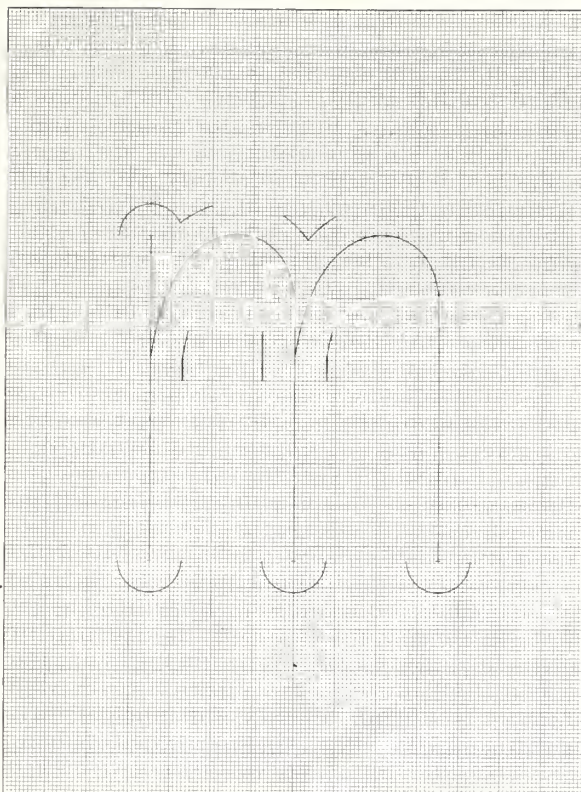


III - 47 Small letter k

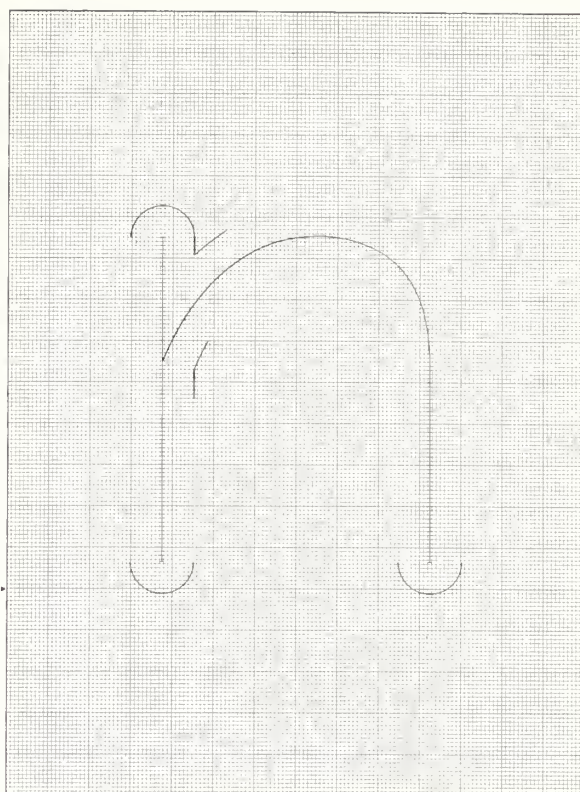


III - 48 Small letter l

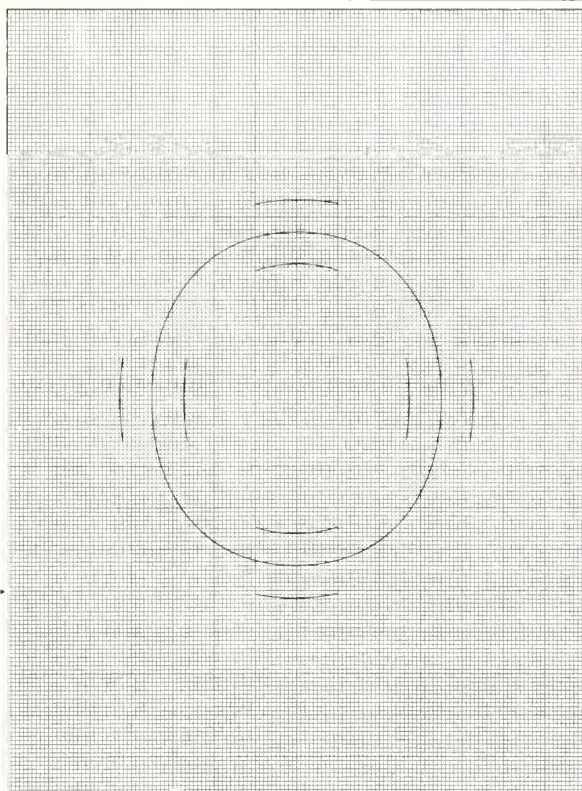




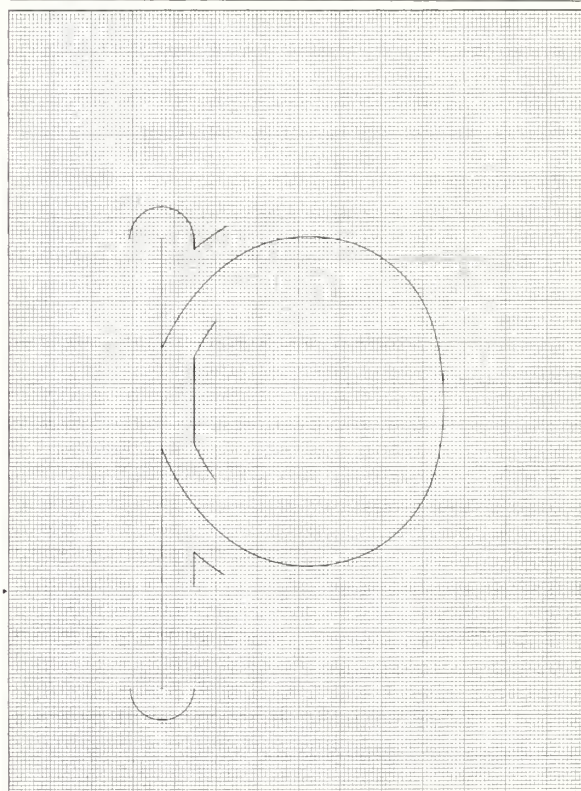
III - 49 Small letter m



III - 50 Small letter n

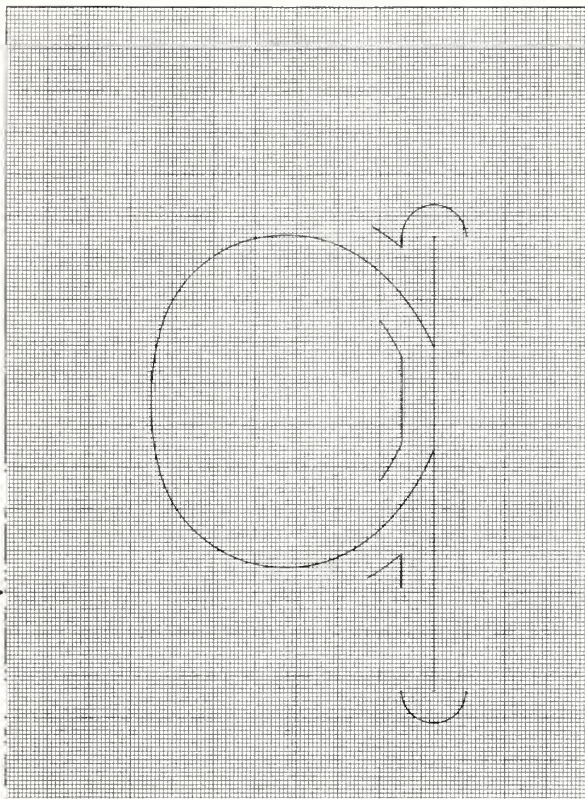


III - 51 Small letter o

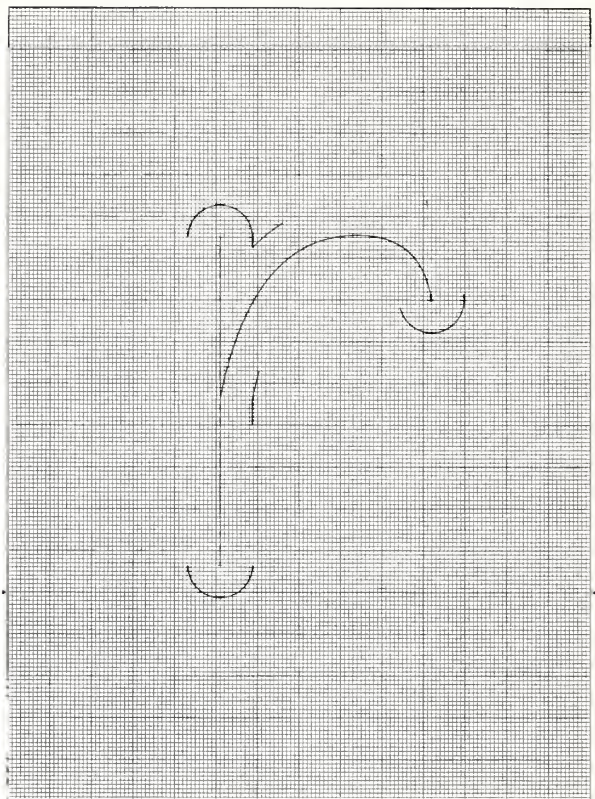


III - 52 Small letter p

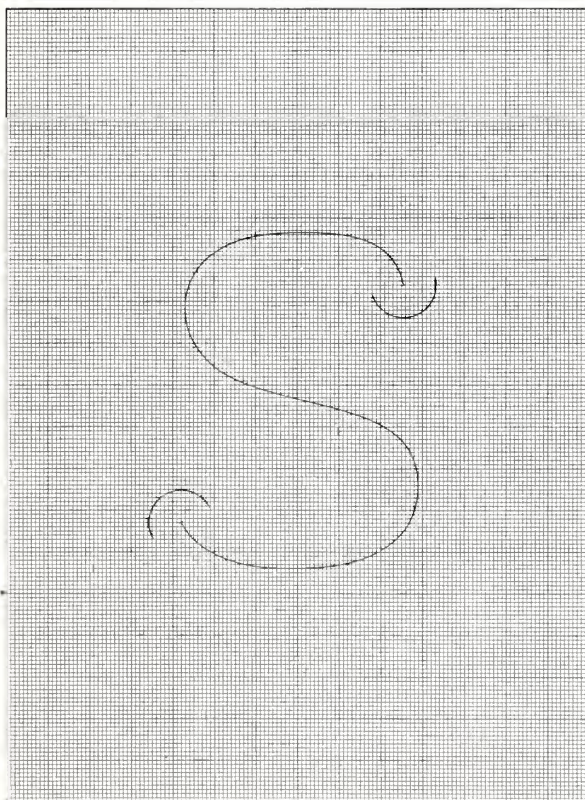




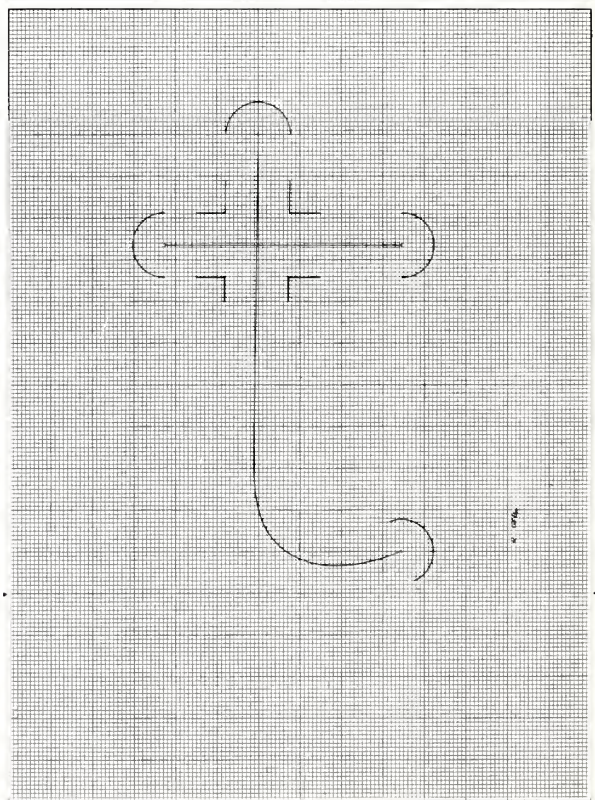
III - 53 Small letter q



III - 54 Small letter r

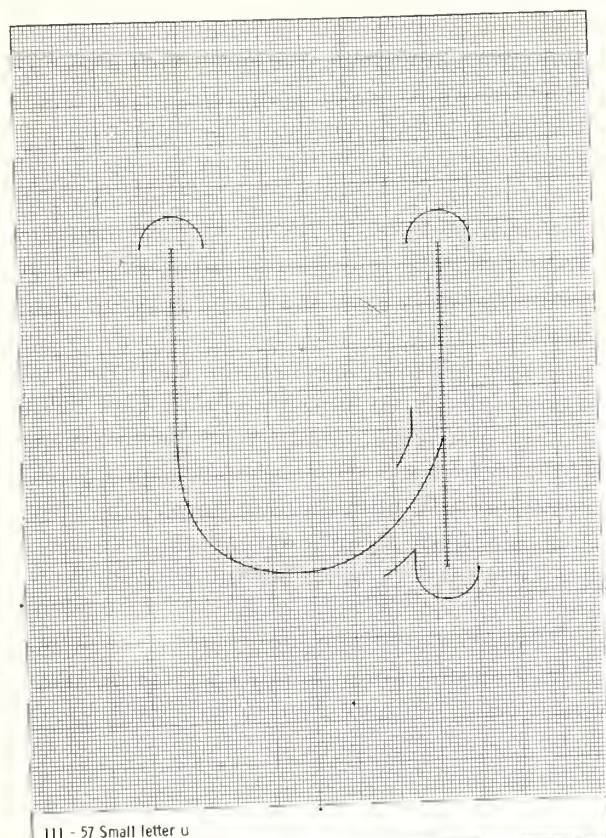


III - 55 Small letter s

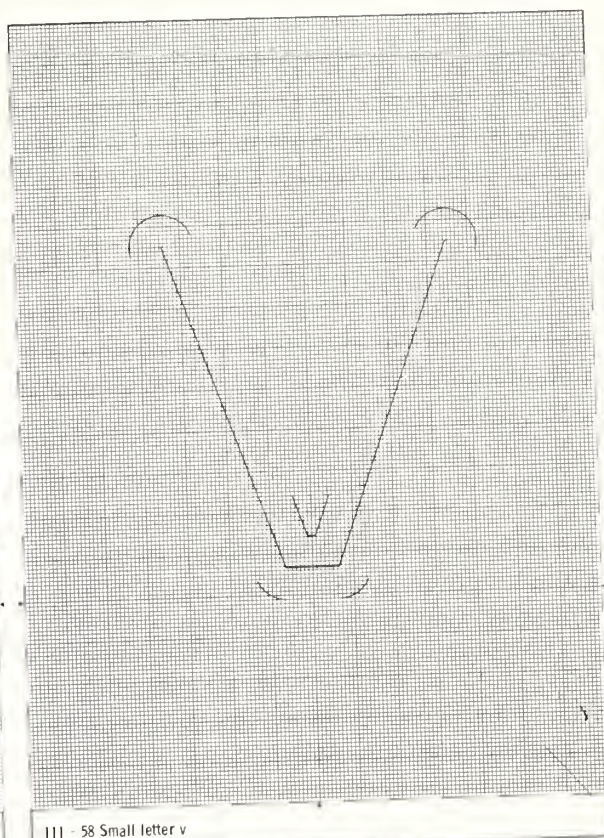


III - 56 Small letter t

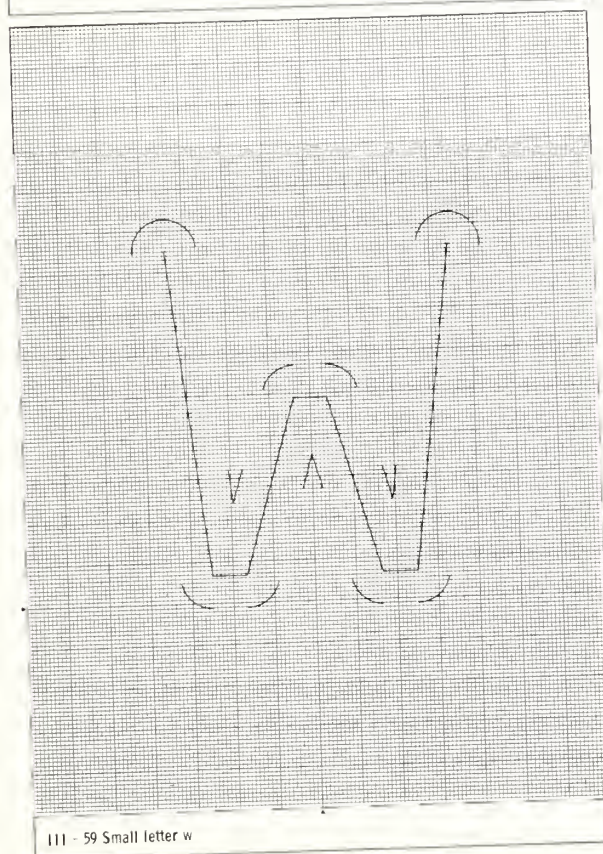




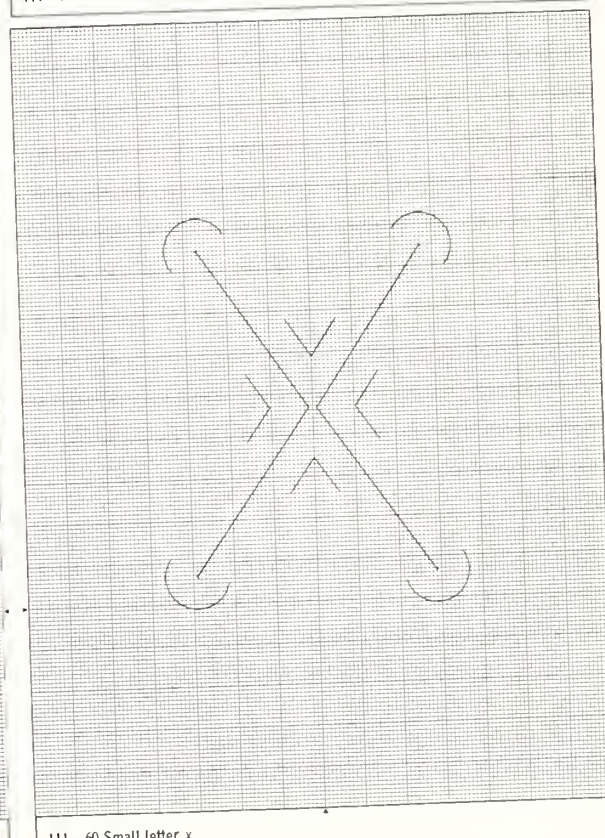
111 - 57 Small letter u



111 - 58 Small letter v

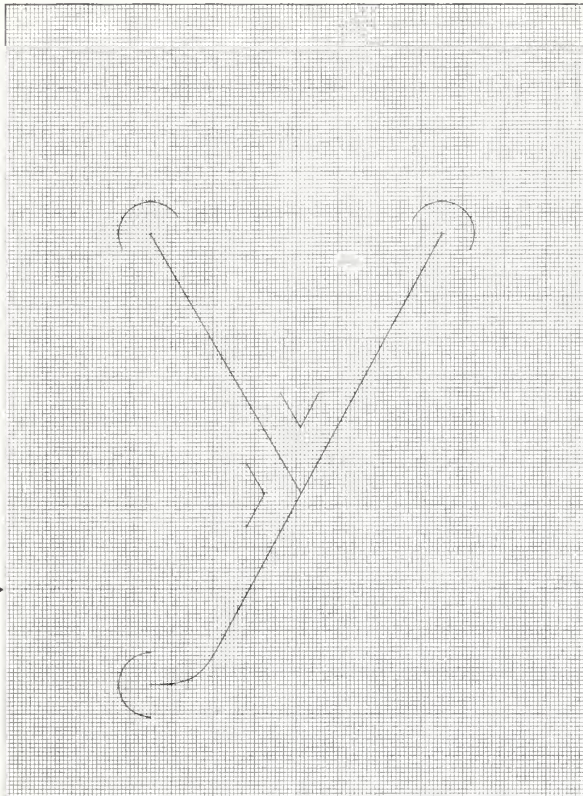


111 - 59 Small letter w

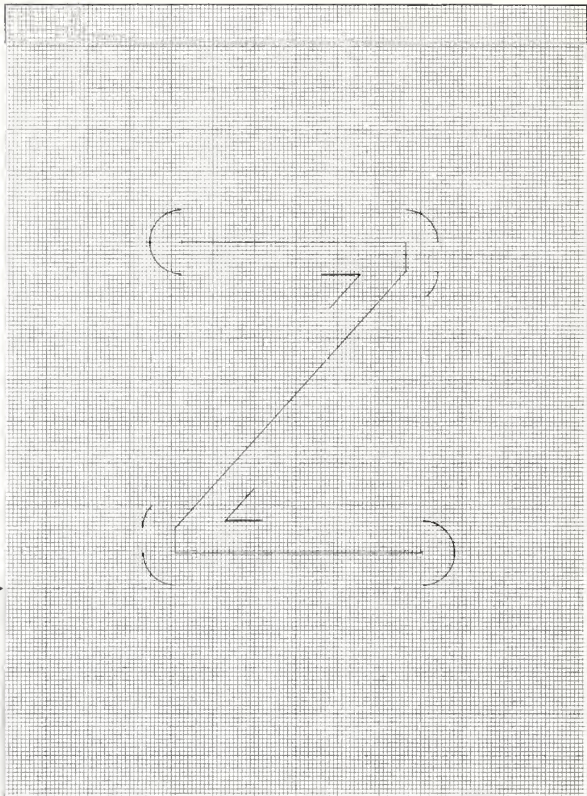


111 - 60 Small letter x

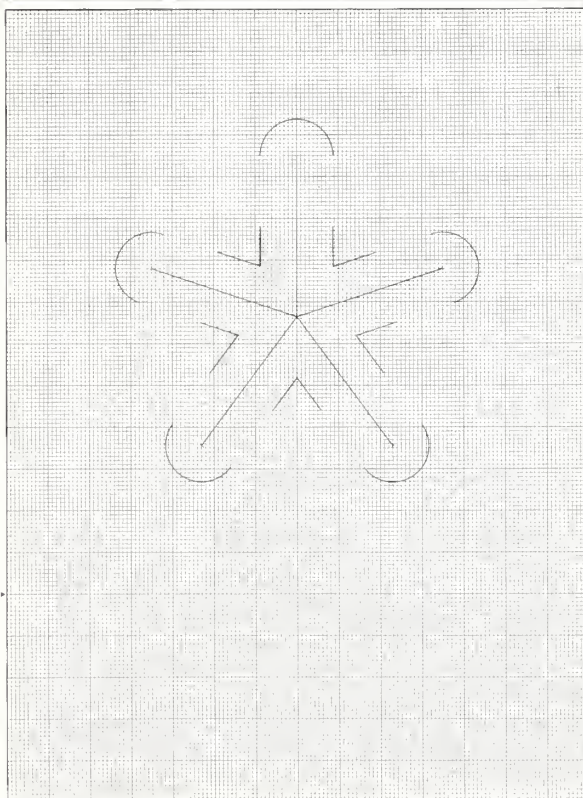




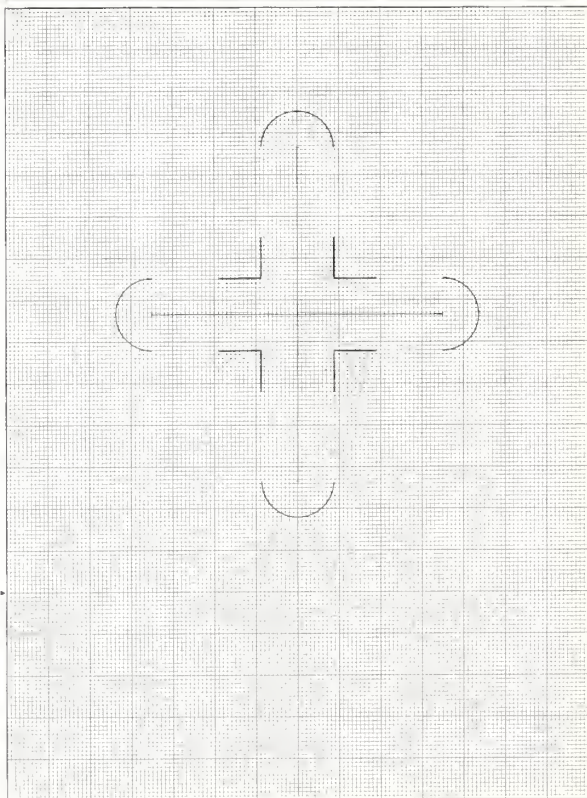
III - 61 Small letter y



III - 62 Small letter z

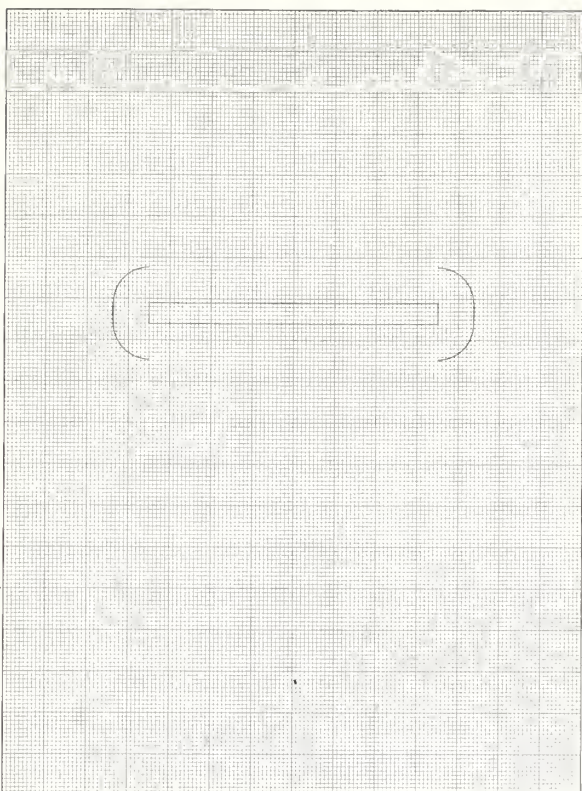


III - 63 ASTERISK

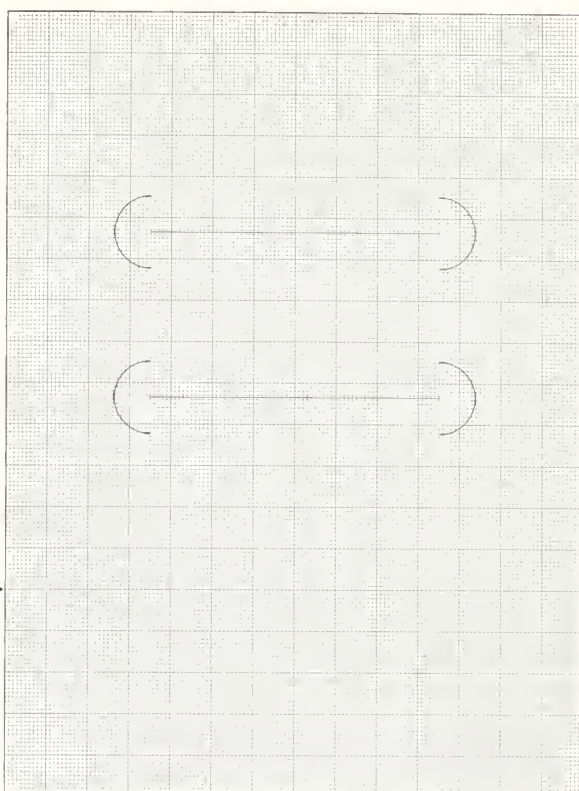


III - 64 PLUS

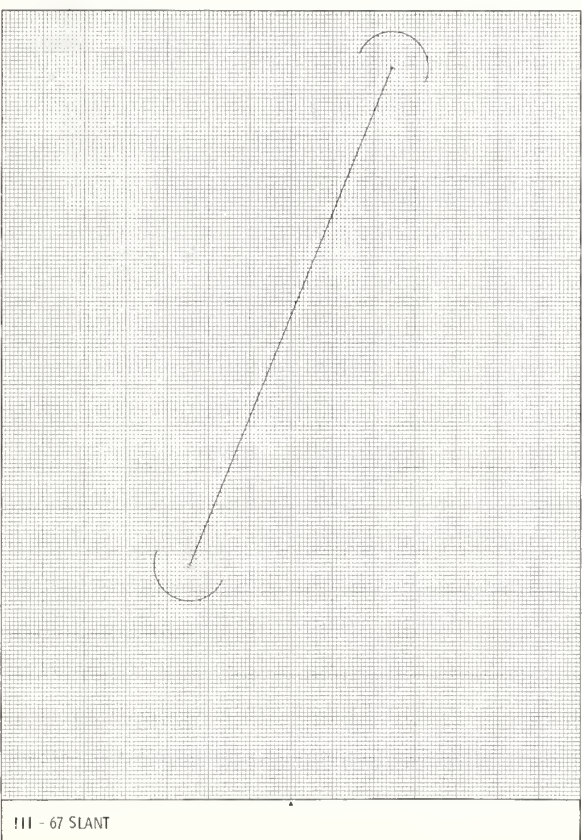




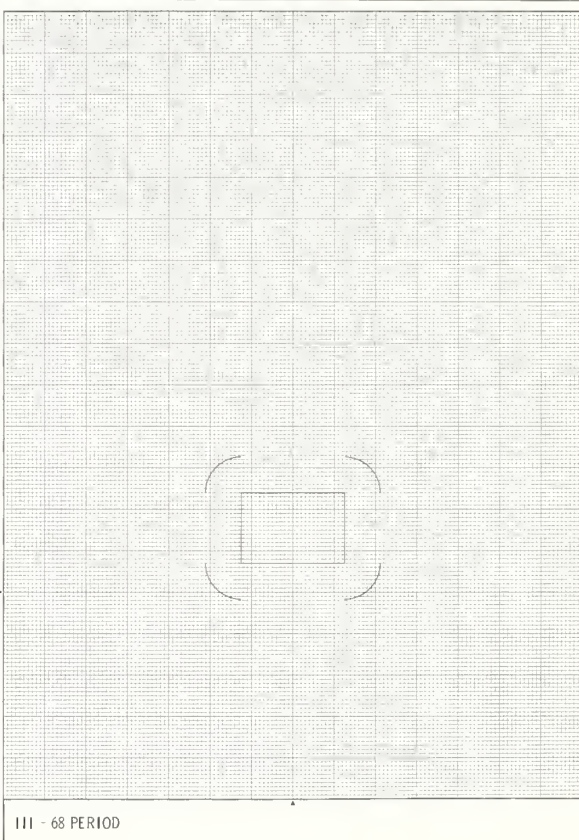
III - 65 HYPHEN



III - 66 EQUALS

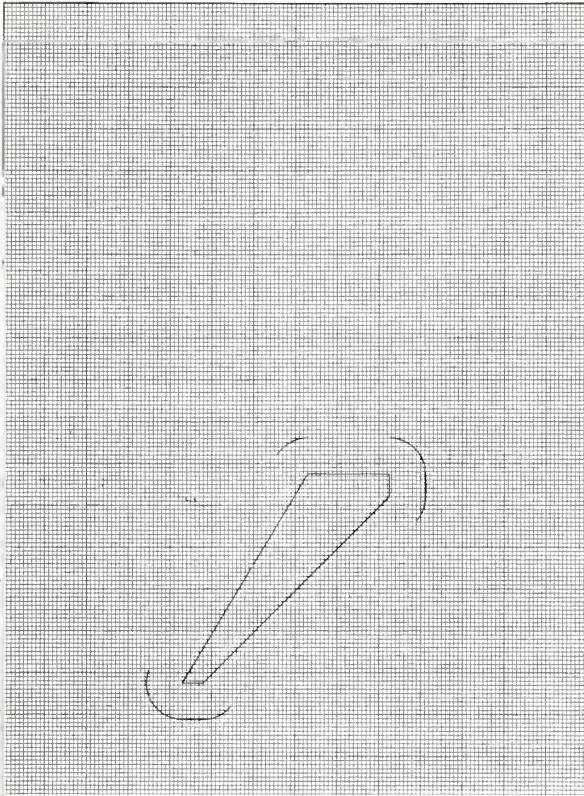


III - 67 SLANT

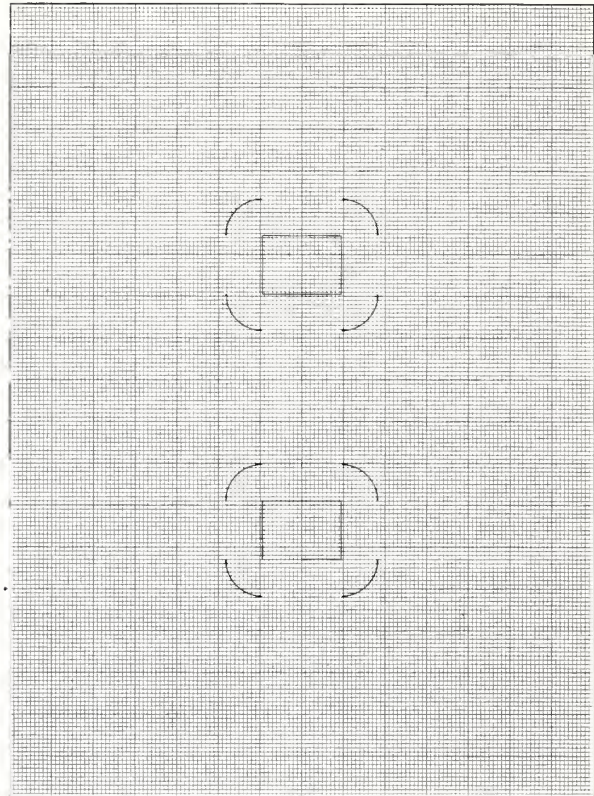


III - 68 PERIOD

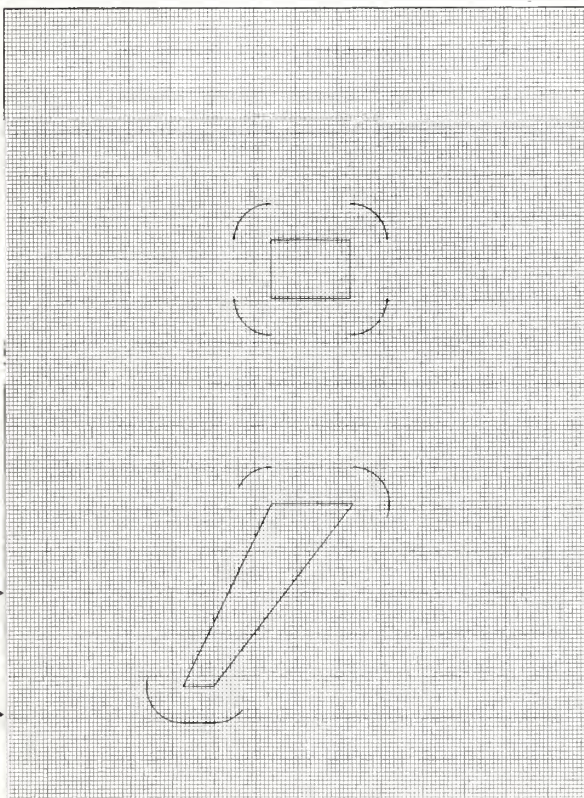




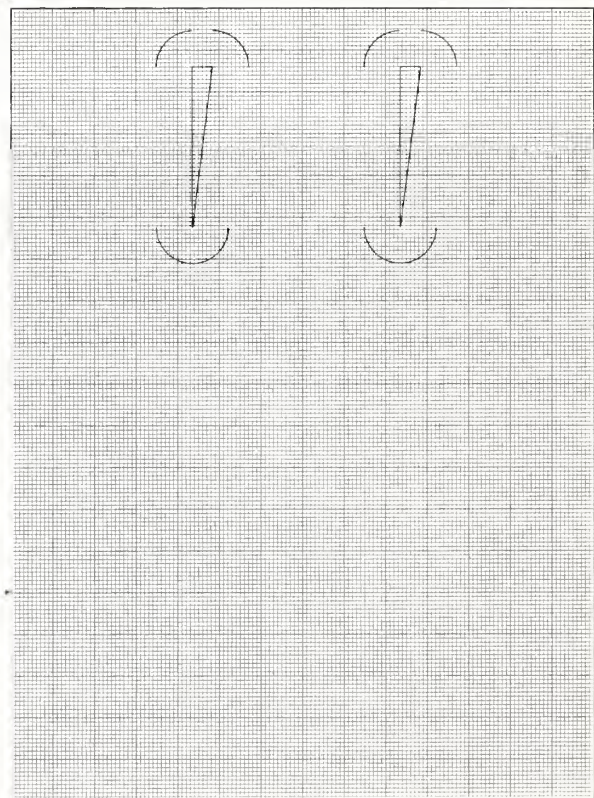
III - 69 COMMA



III - 70 COLON

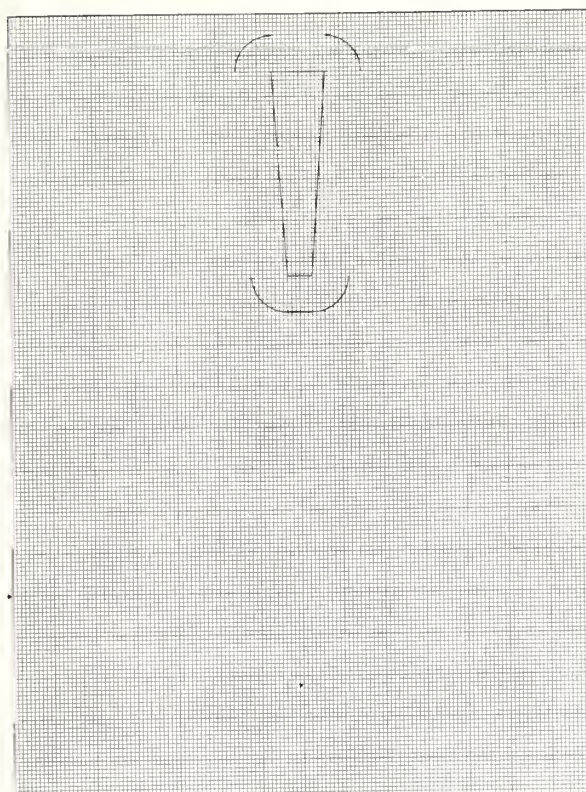


III - 71 SEMICOLON

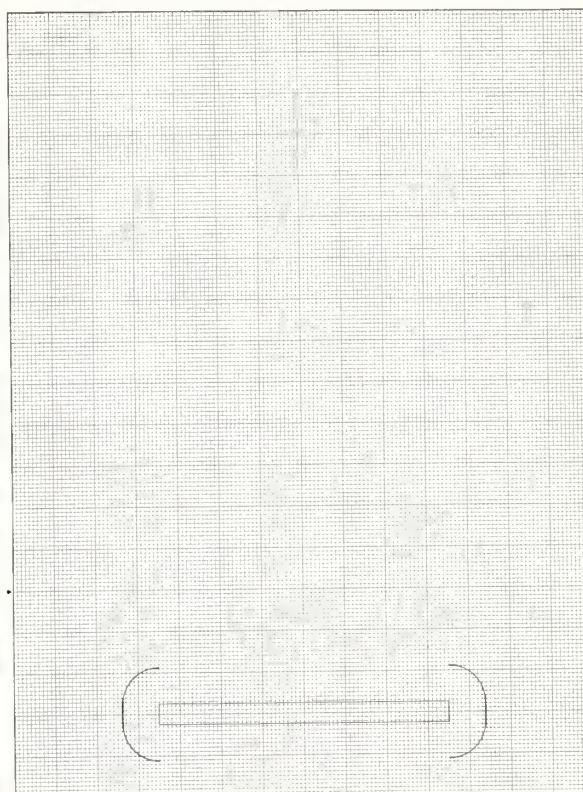


III - 72 QUOTATION MARKS

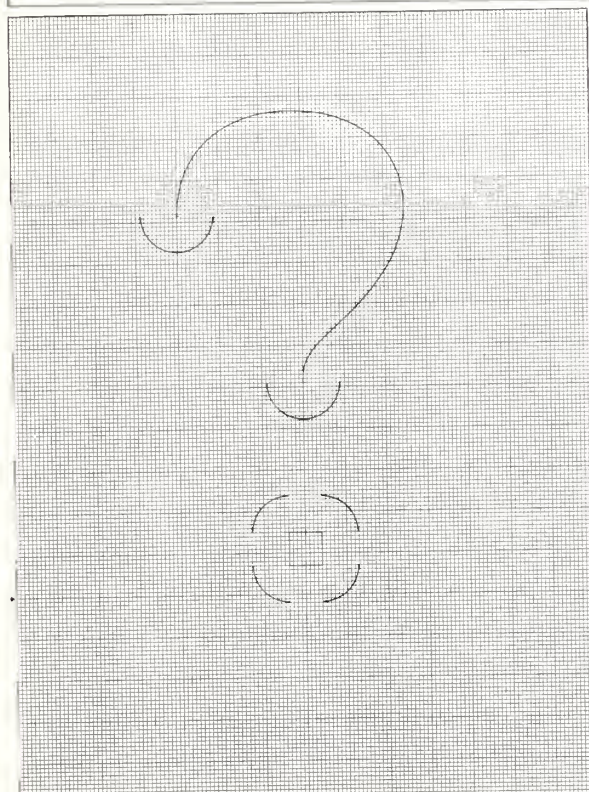




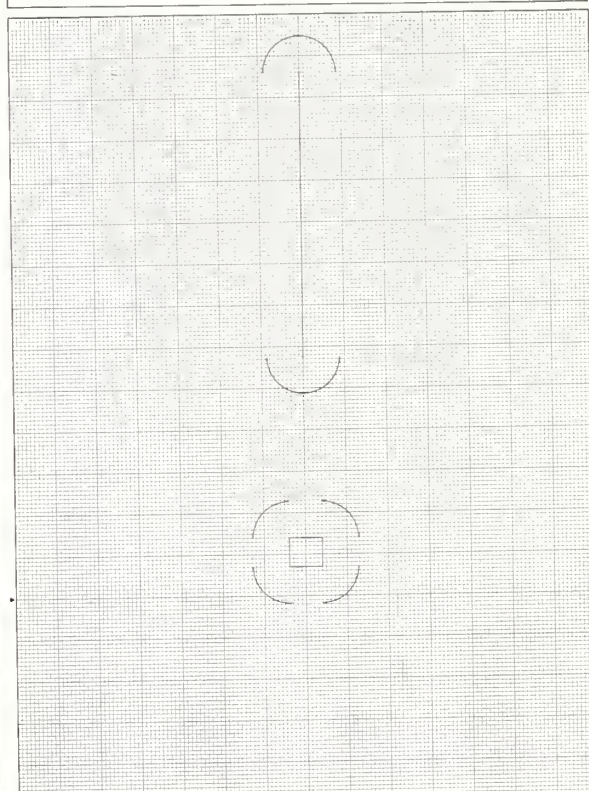
III - 73 APOSTROPHE



III - 74 DISCONTINUOUS UNDERLINE

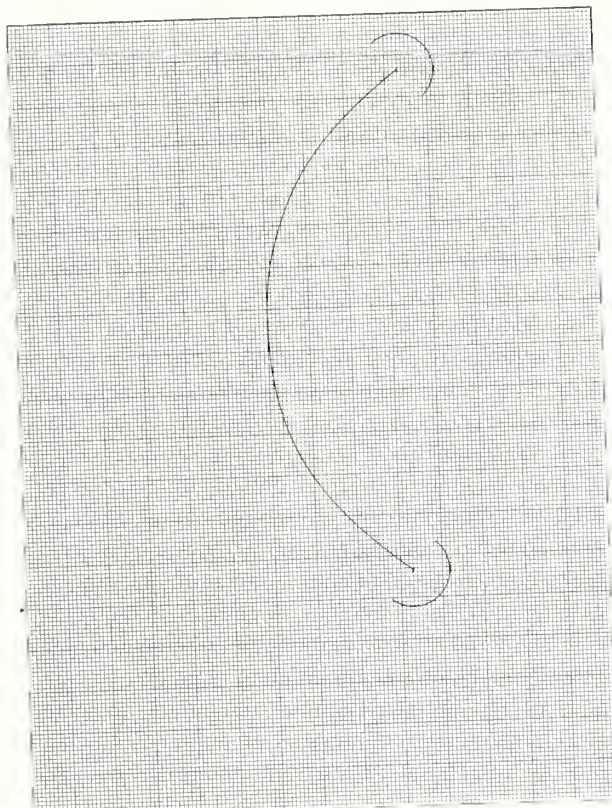


III - 75 QUESTION MARK

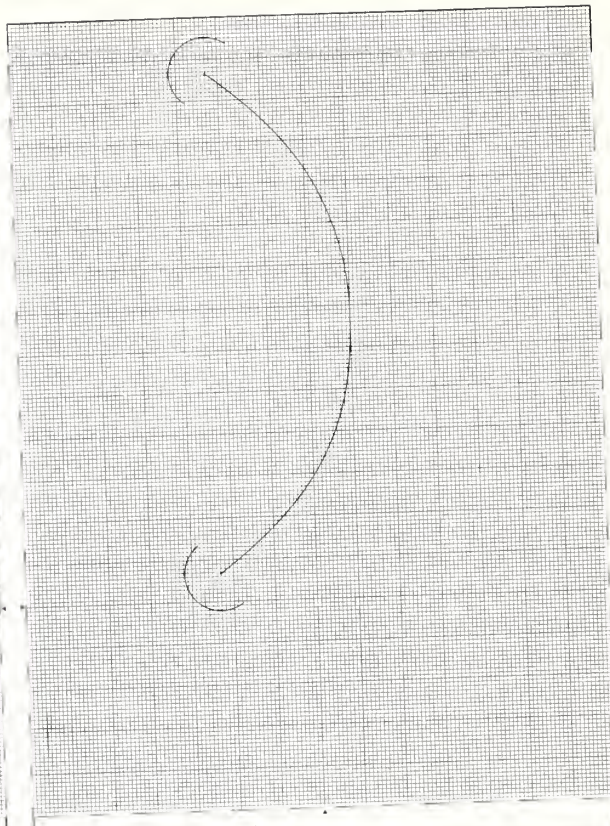


III - 76 EXCLAMATION MARK

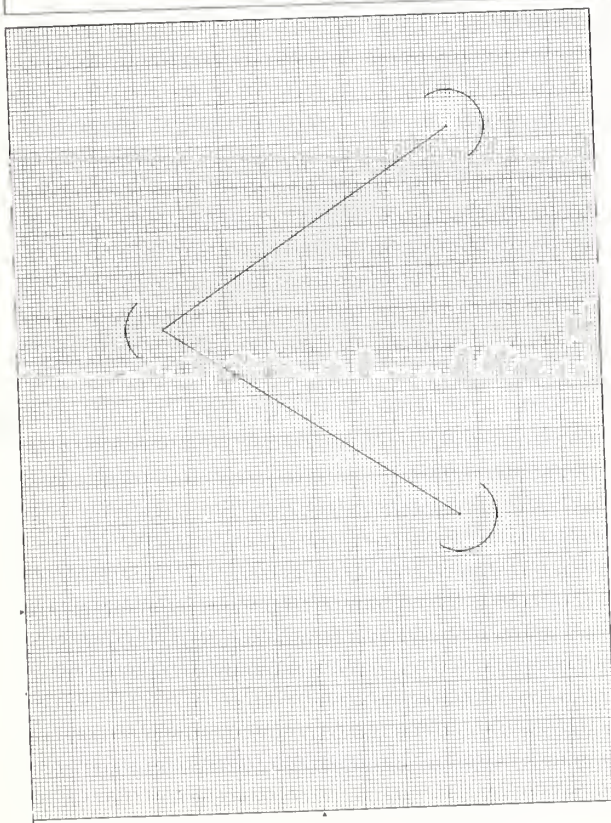




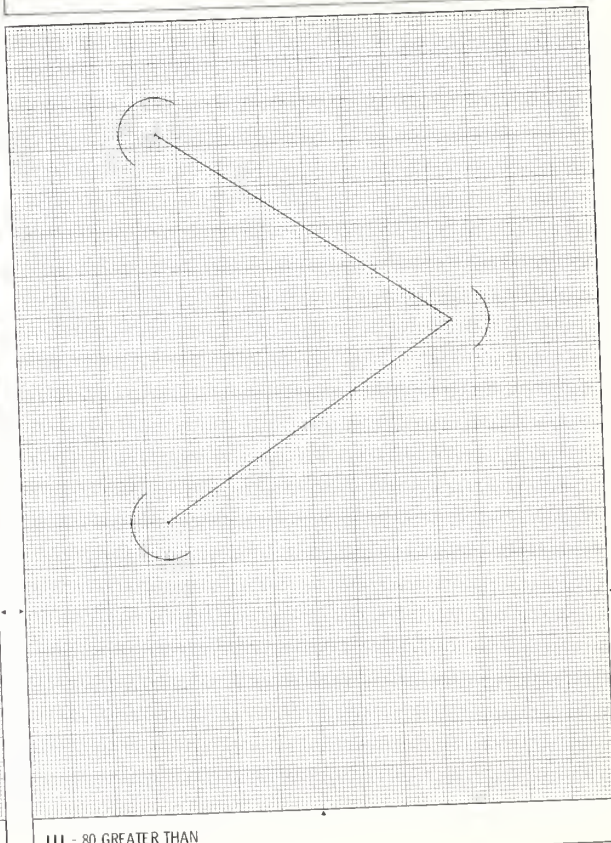
III - 77 OPENING PARENTHESIS



III - 78 CLOSING PARENTHESIS

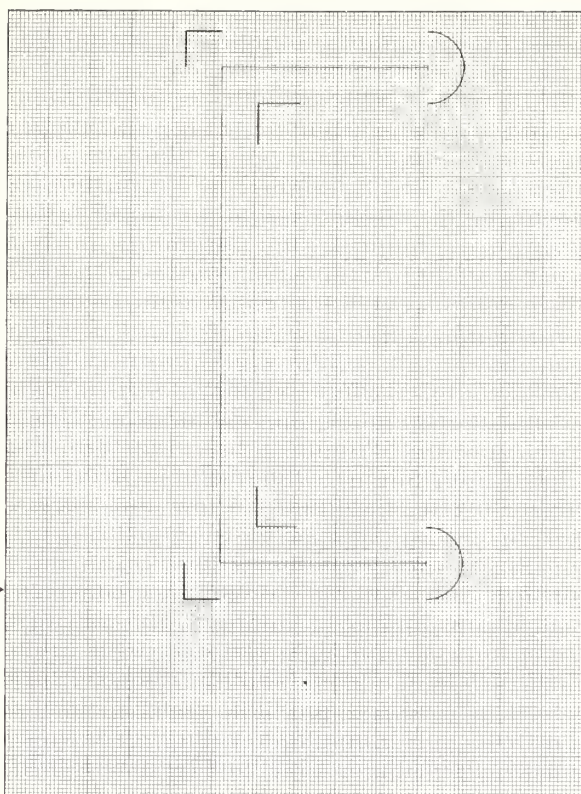


III - 79 LESS THAN

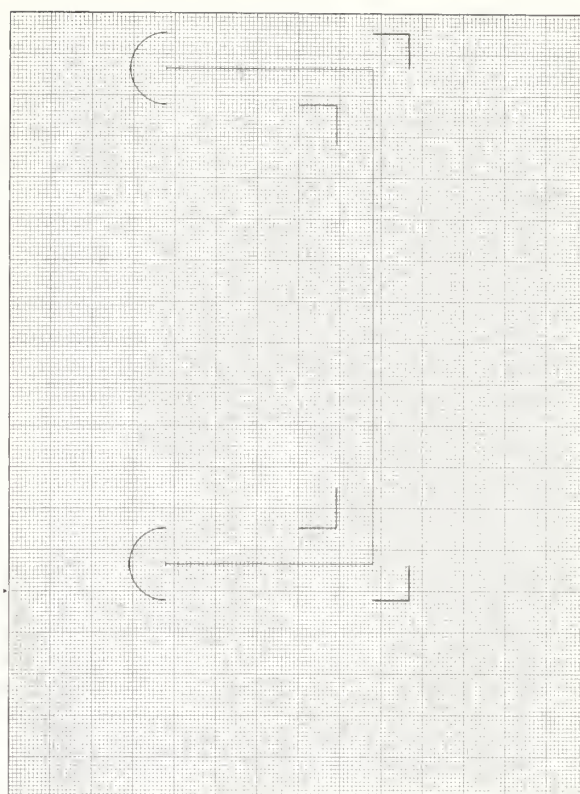


III - 80 GREATER THAN

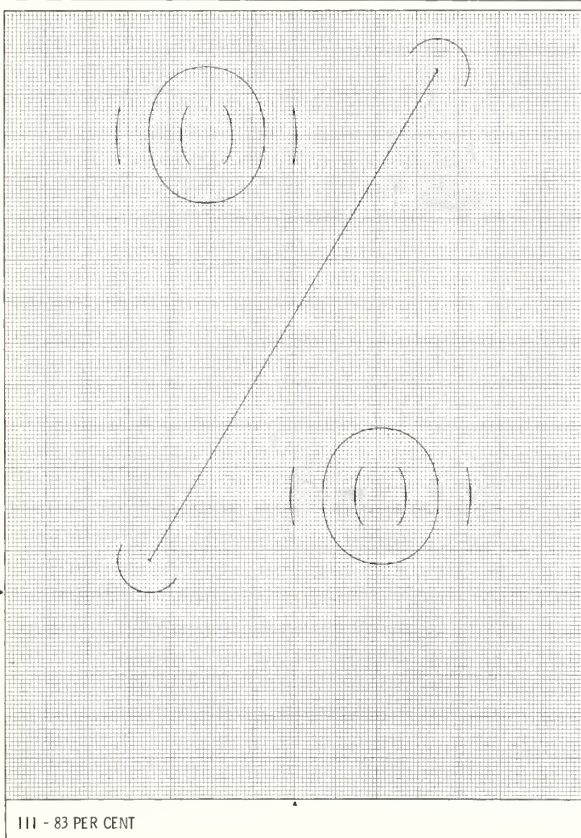




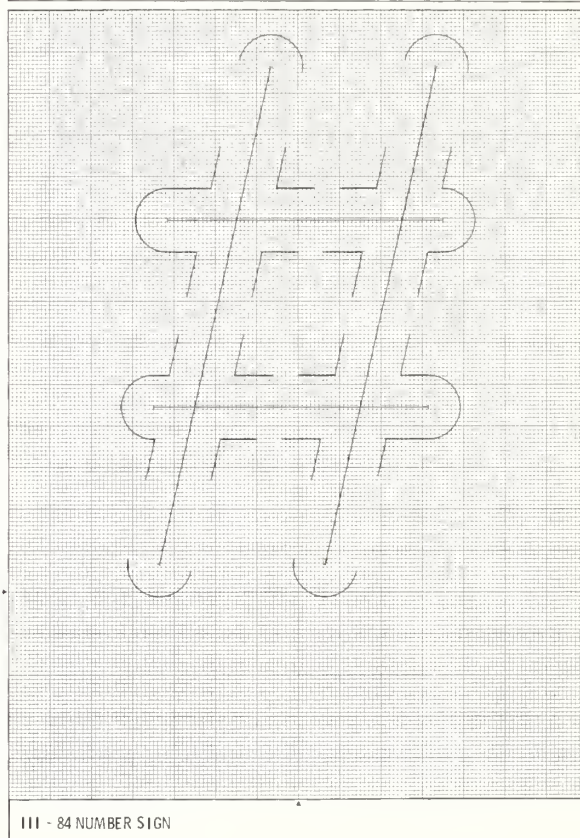
III - 81 OPENING BRACKET



III - 82 CLOSING BRACKET

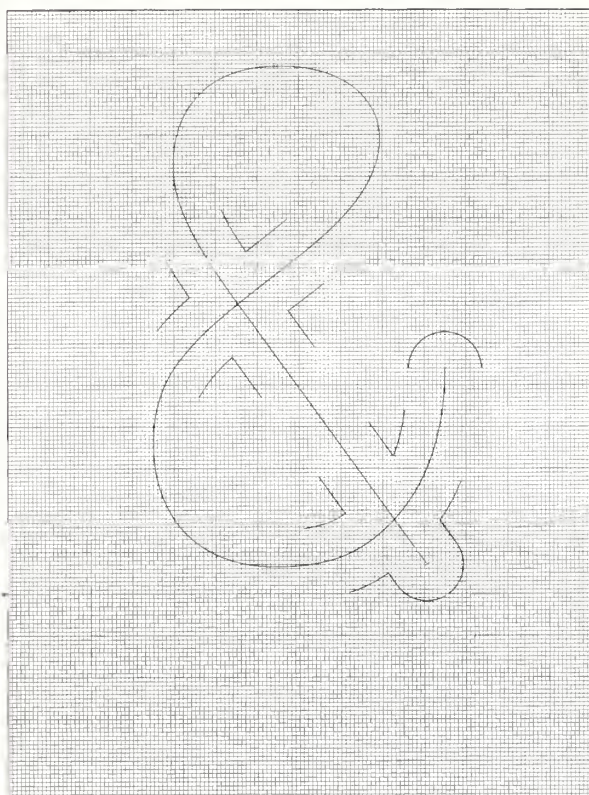


III - 83 PER CENT

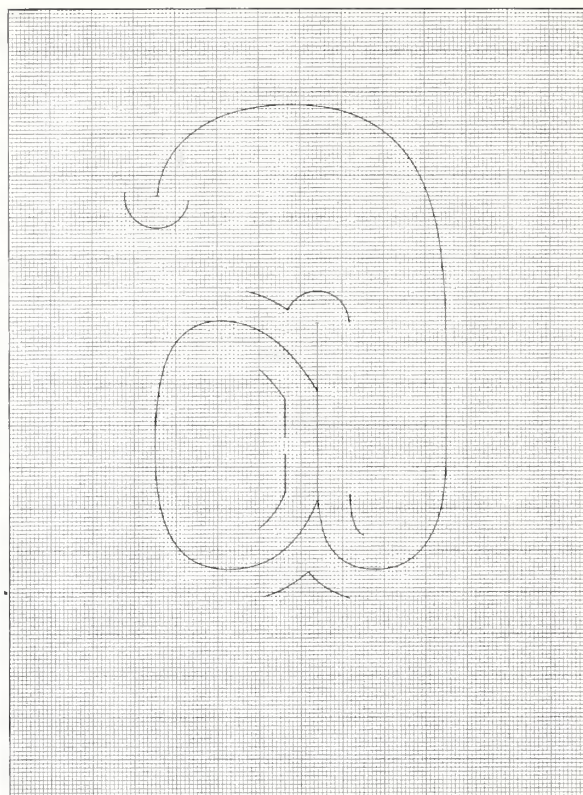


III - 84 NUMBER SIGN

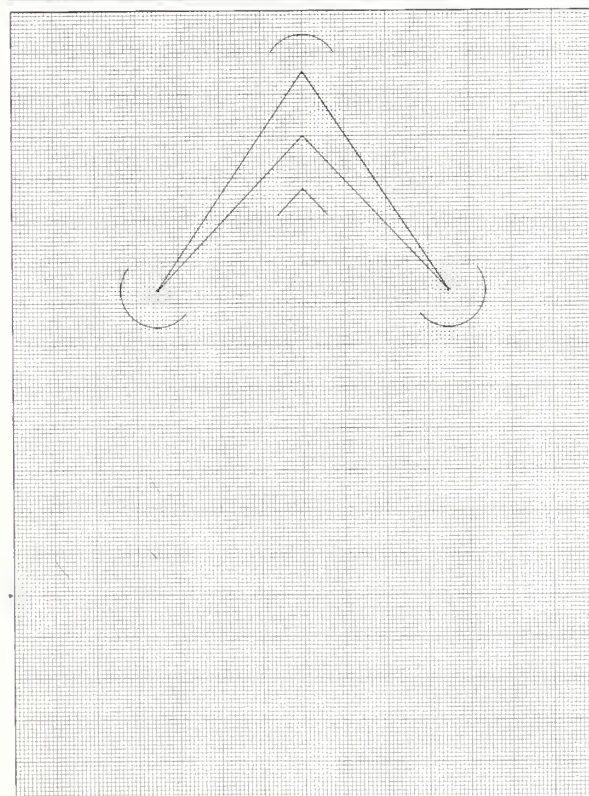




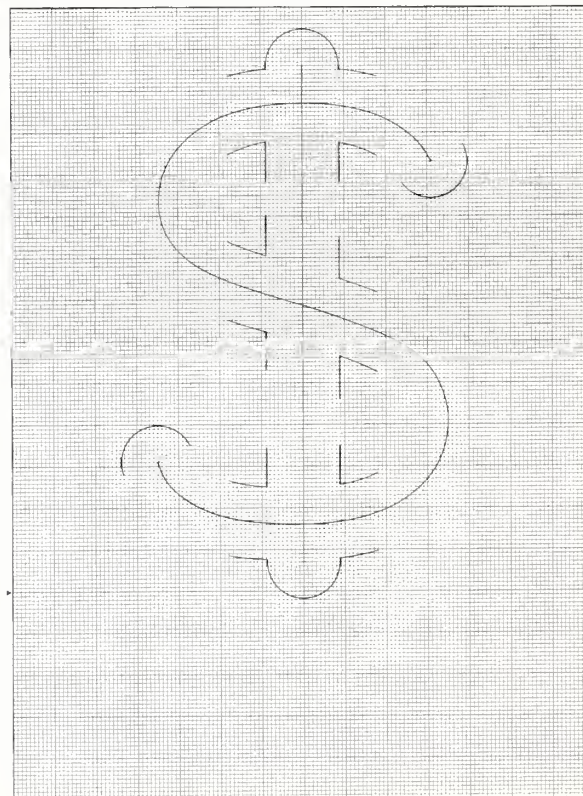
III - 85 AMPERSAND



III - 86 COMMERCIAL AT



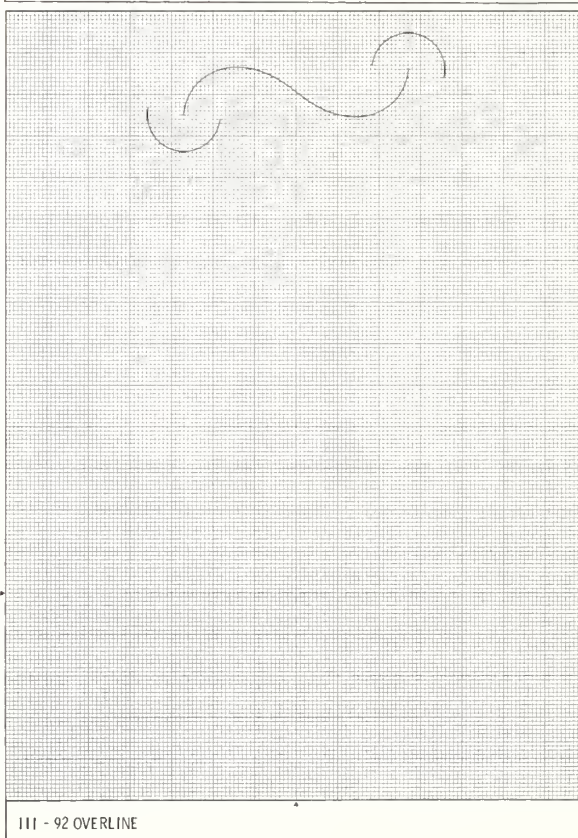
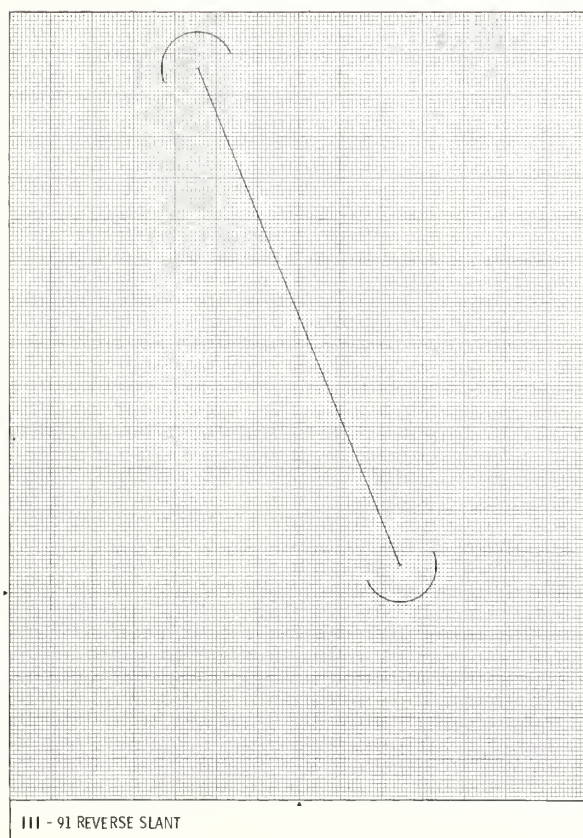
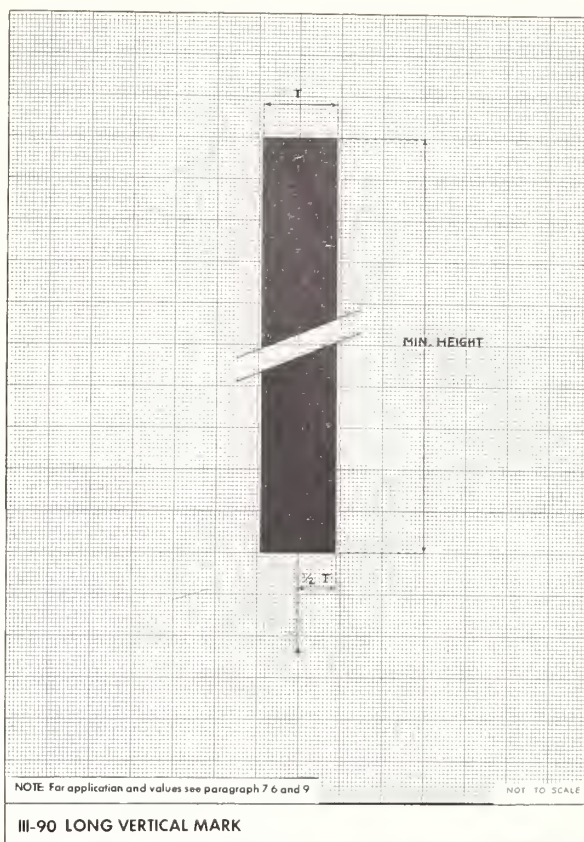
III - 87 UPWARD ARROW HEAD



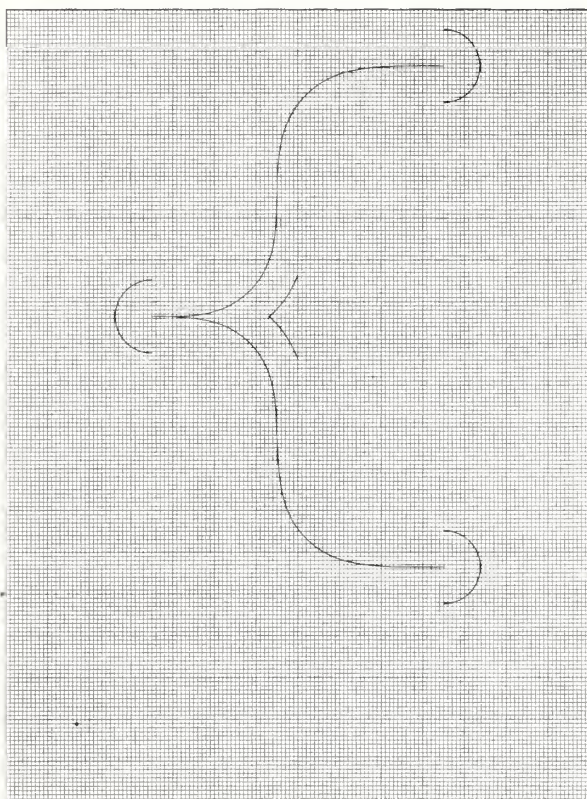
III - 88 DOLLAR SIGN



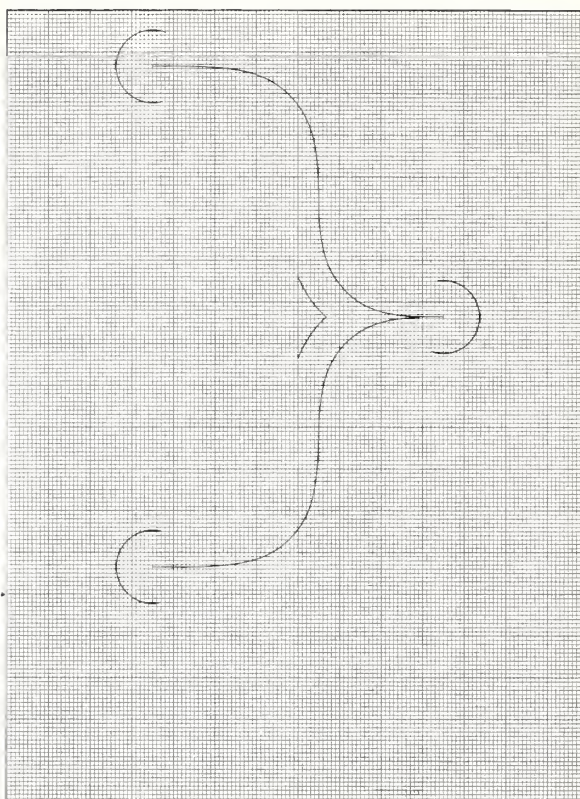
III-89 NOT USED



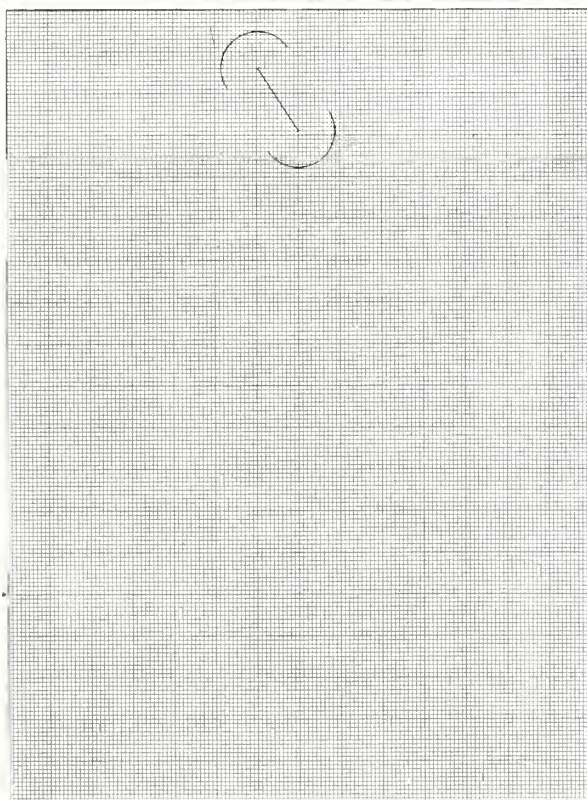




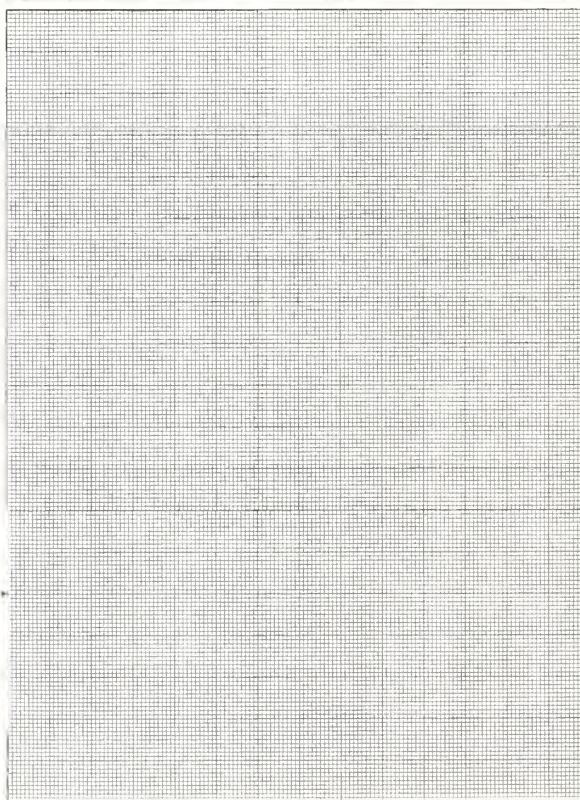
III - 93 OPENING BRACE



III - 94 CLOSING BRACE

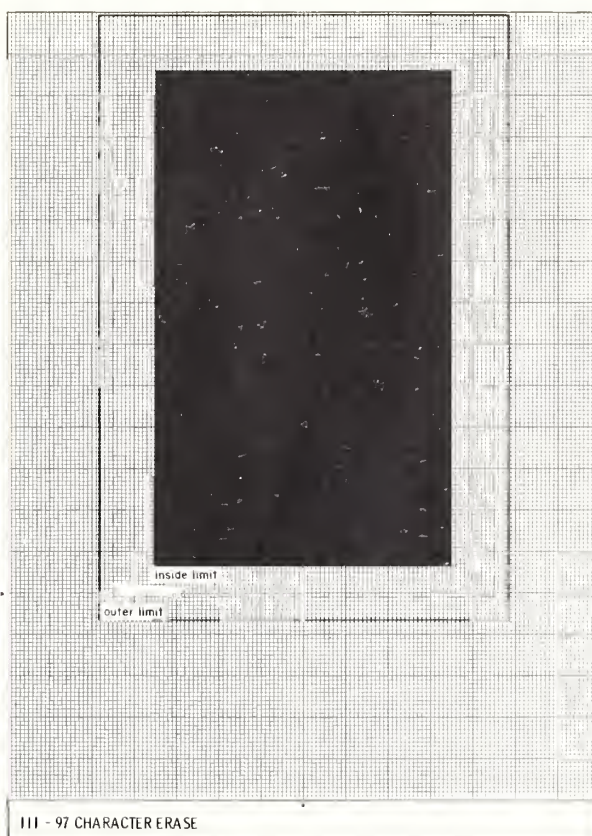


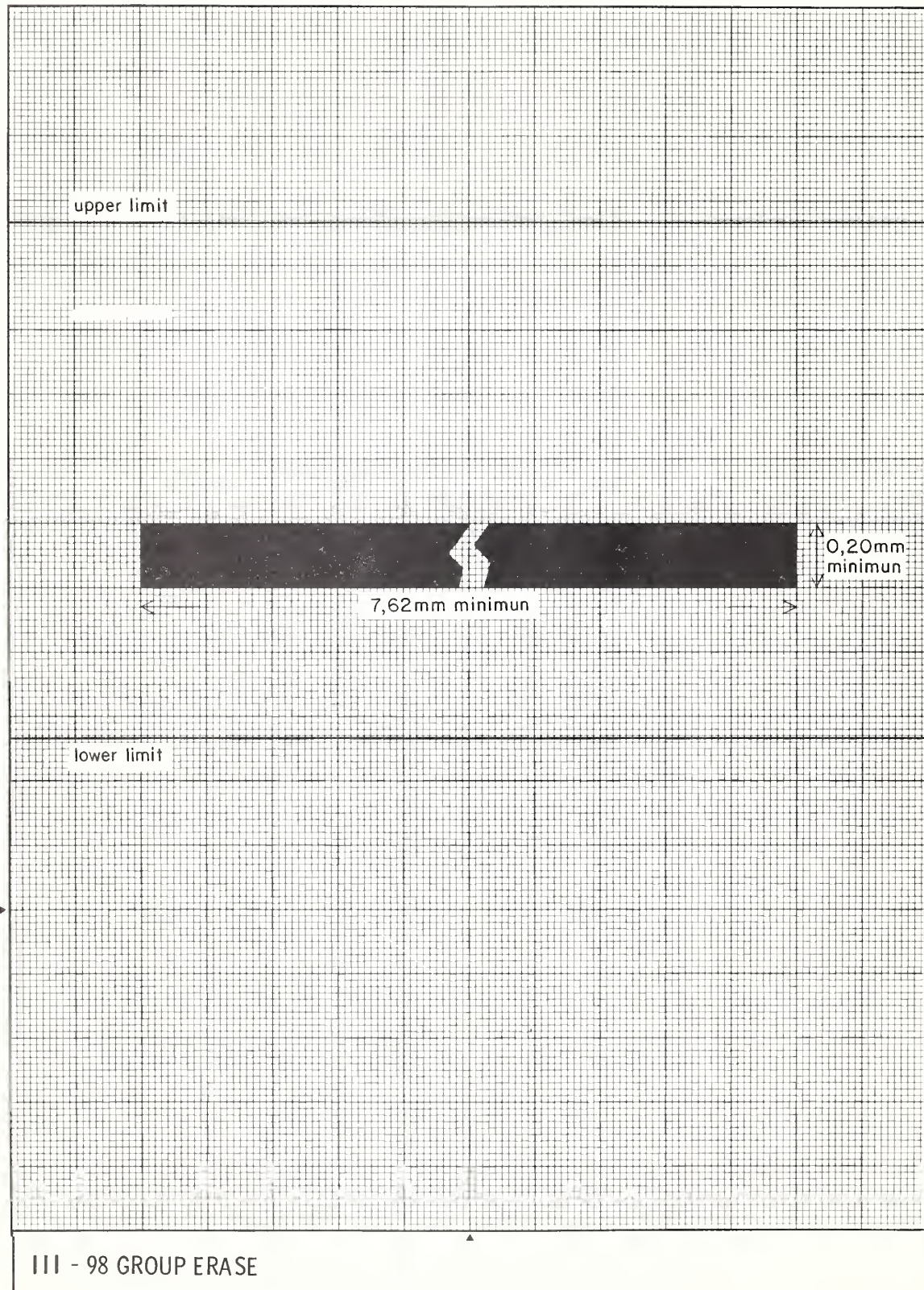
III - 95 GRAVE ACCENT



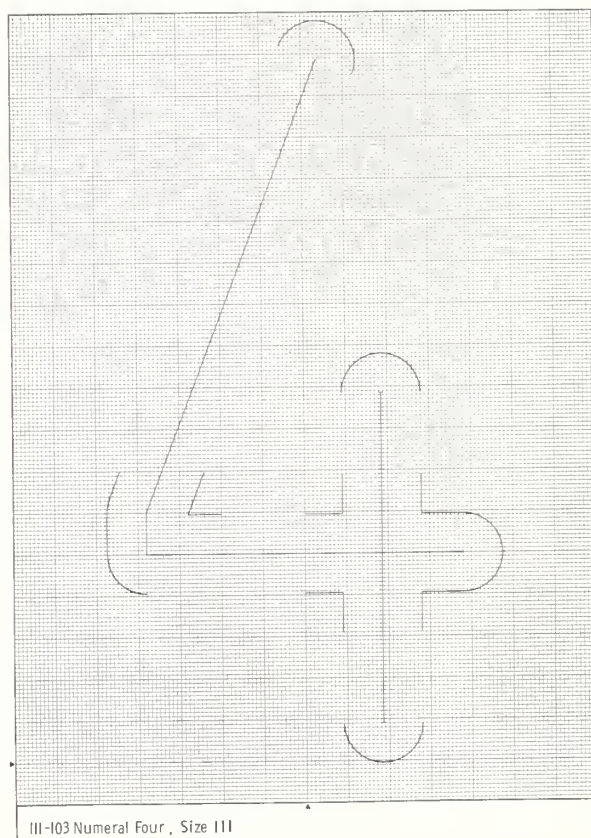
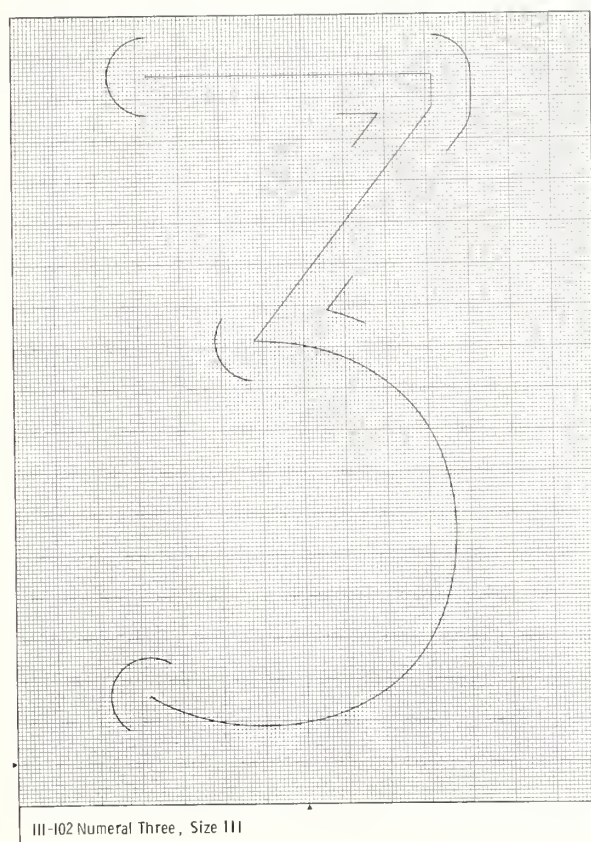
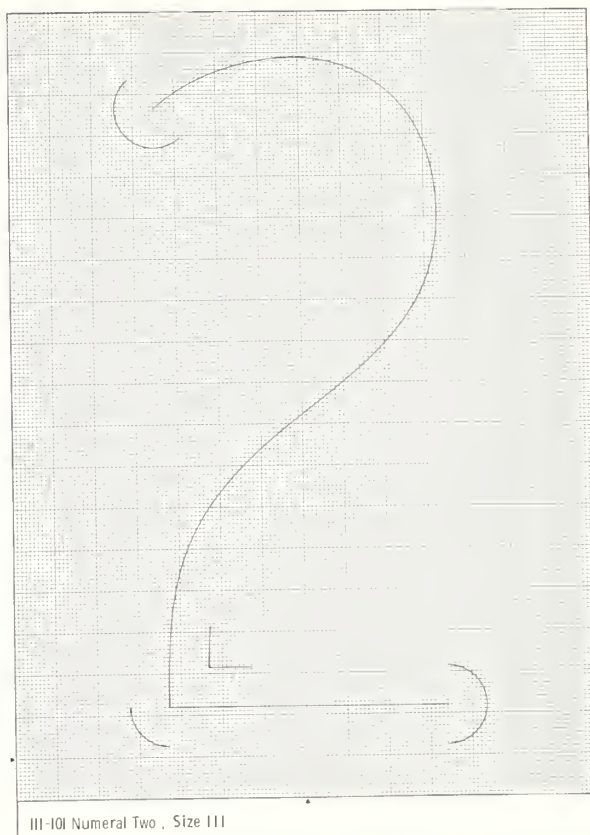
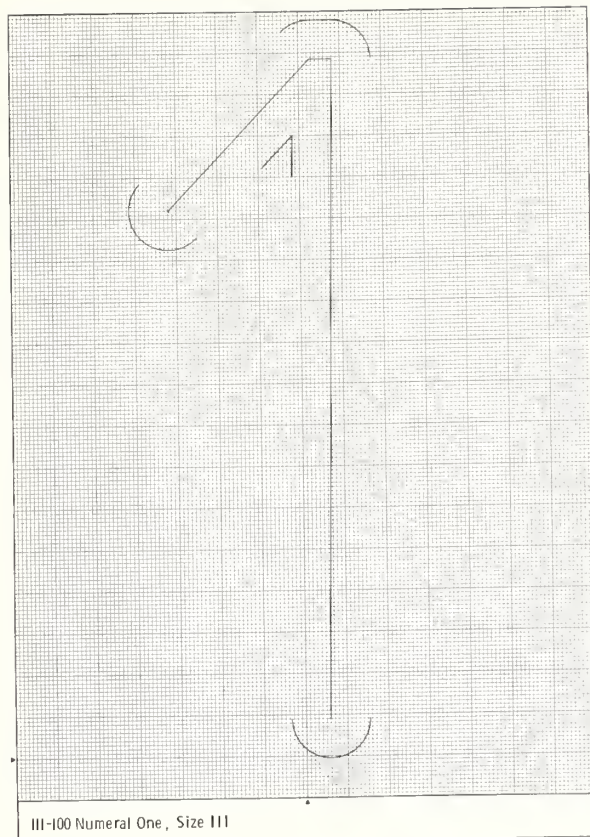
III - 96 Character SPACE



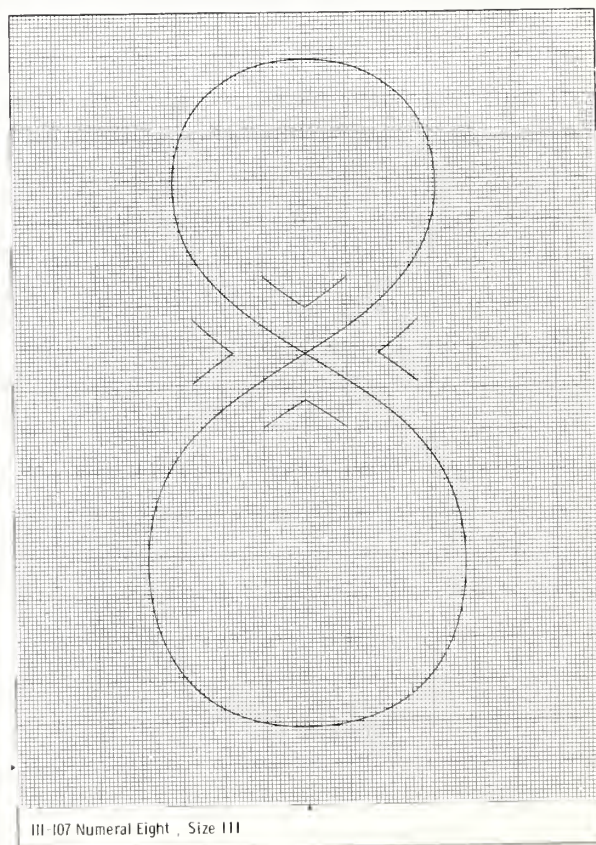
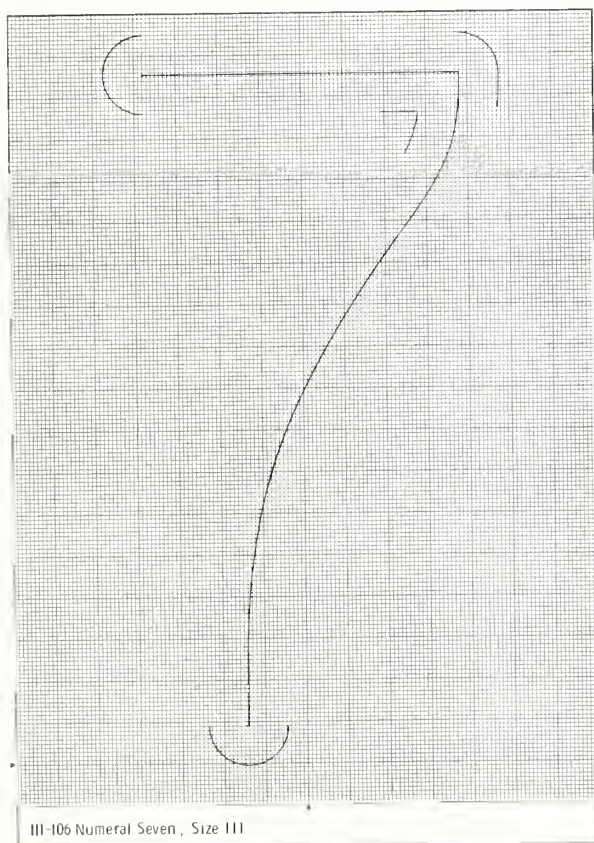
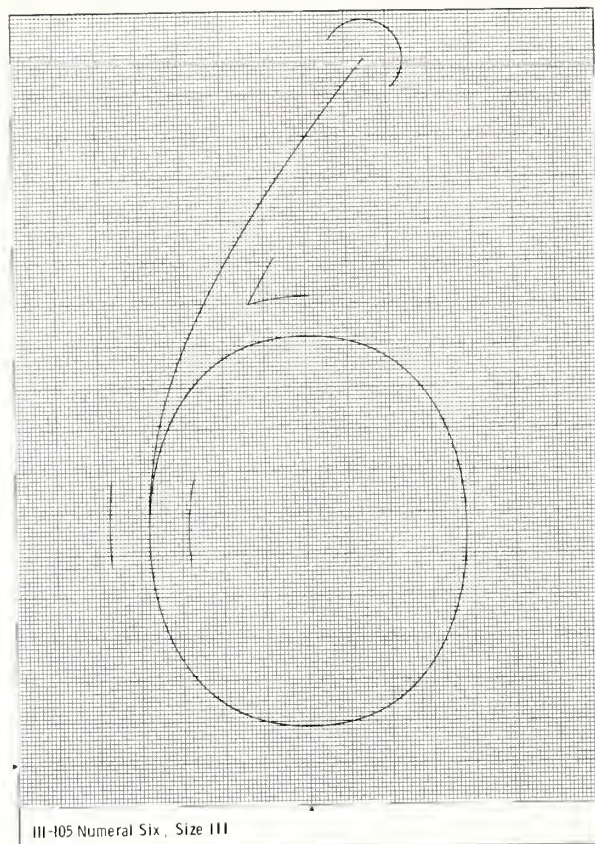
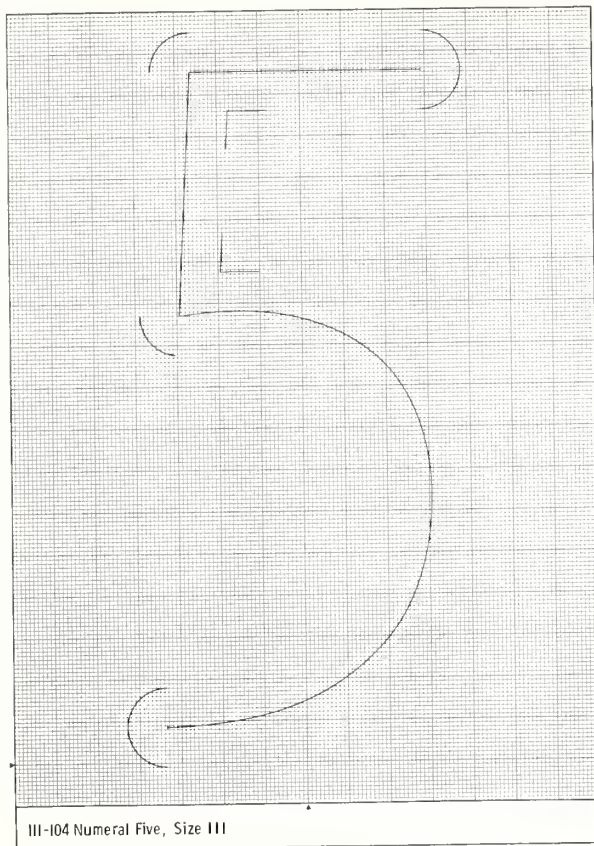




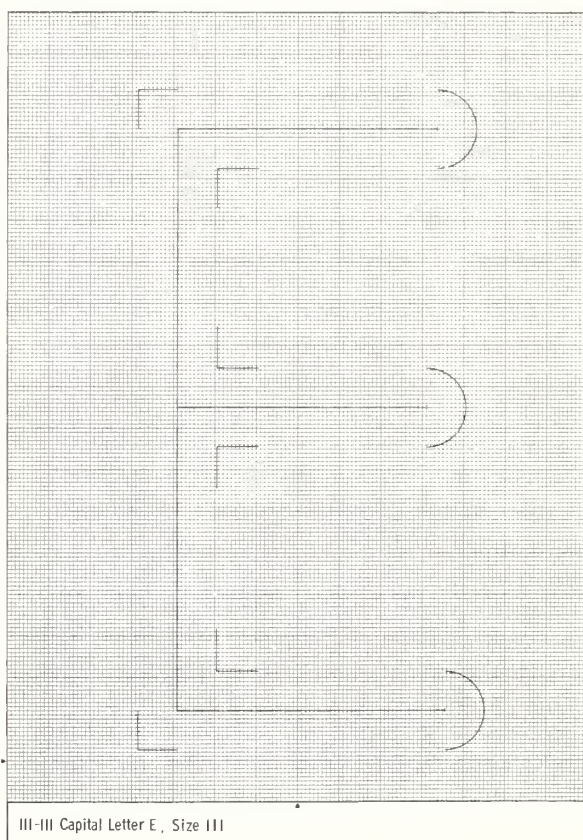
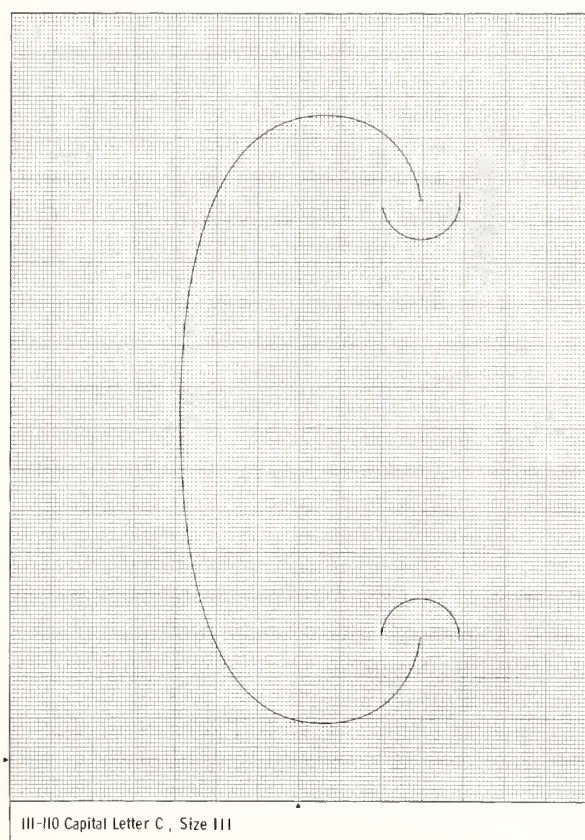
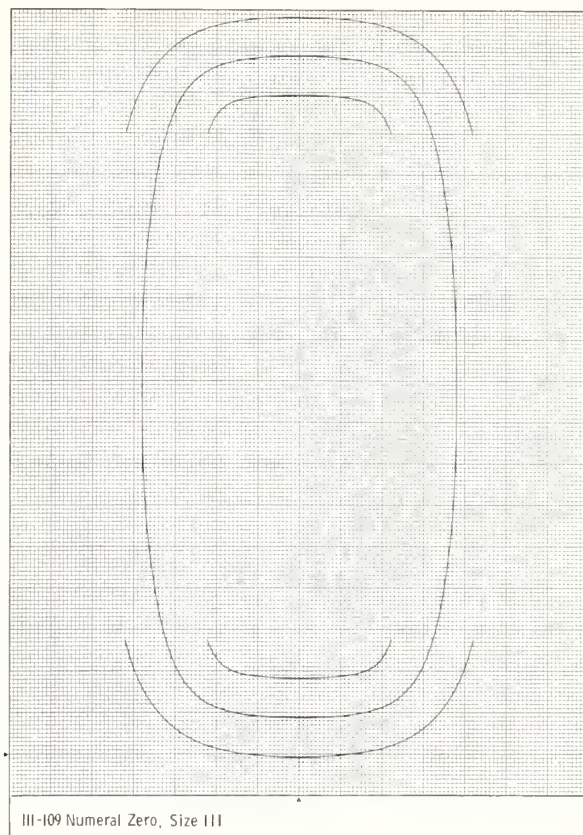
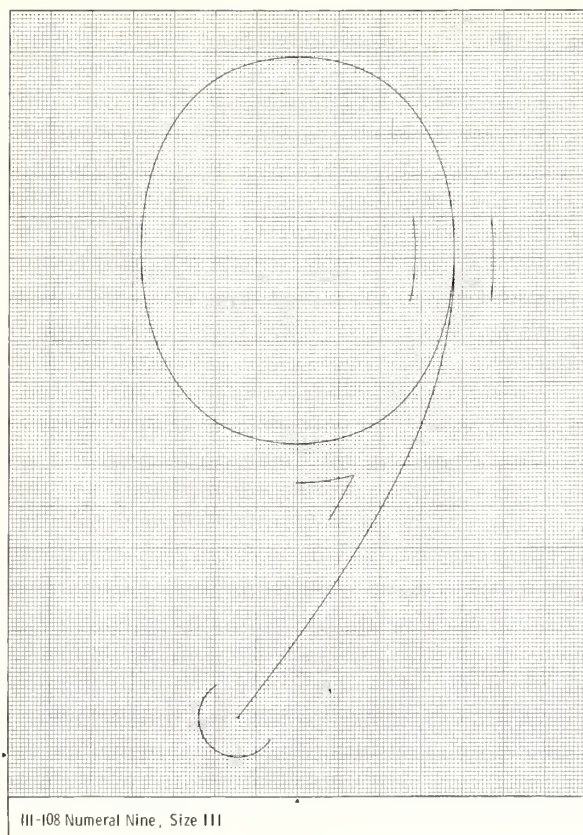




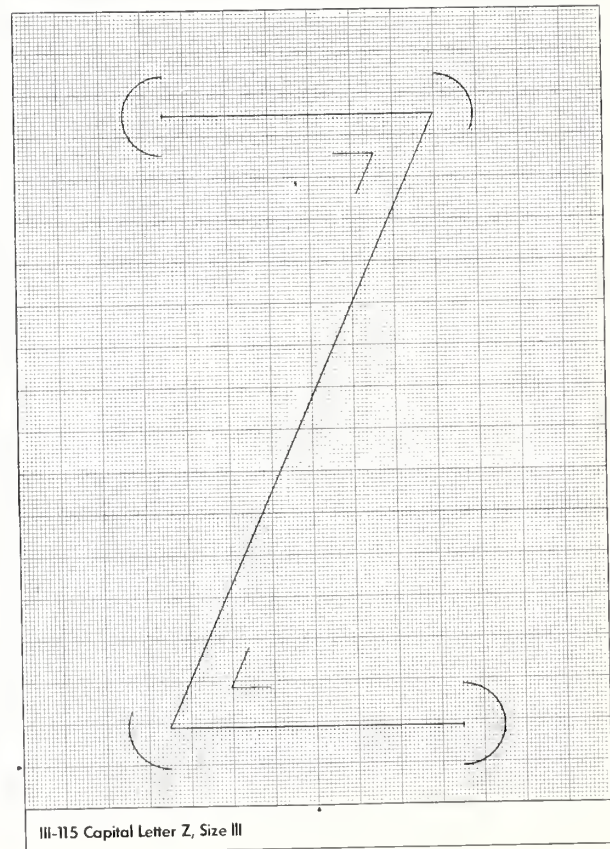
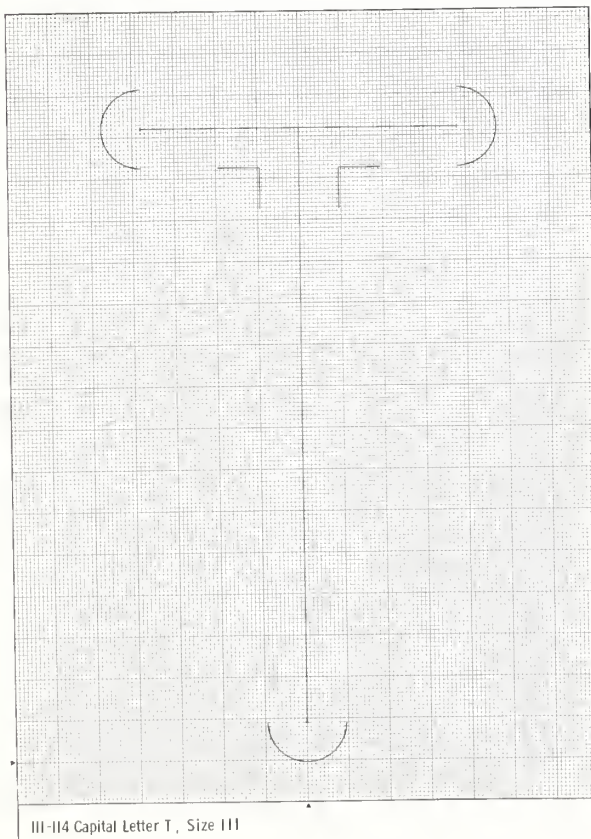
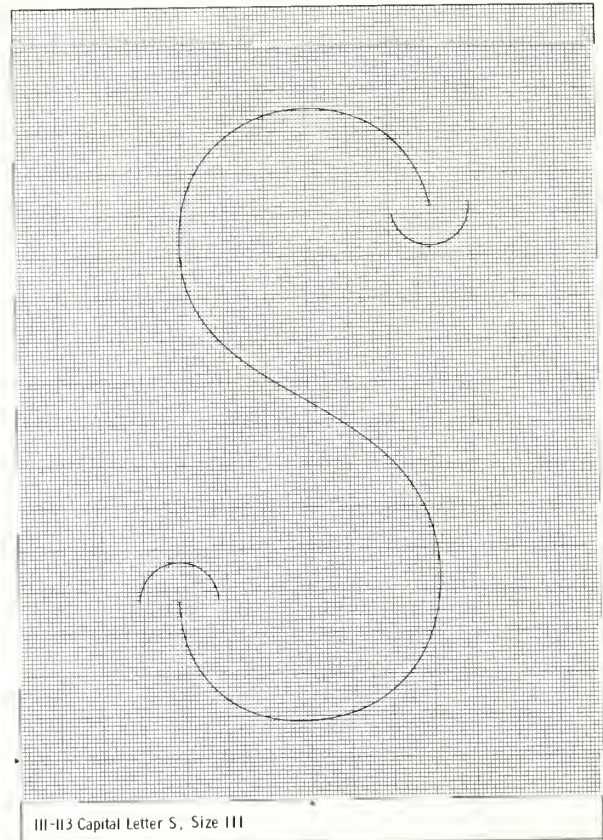
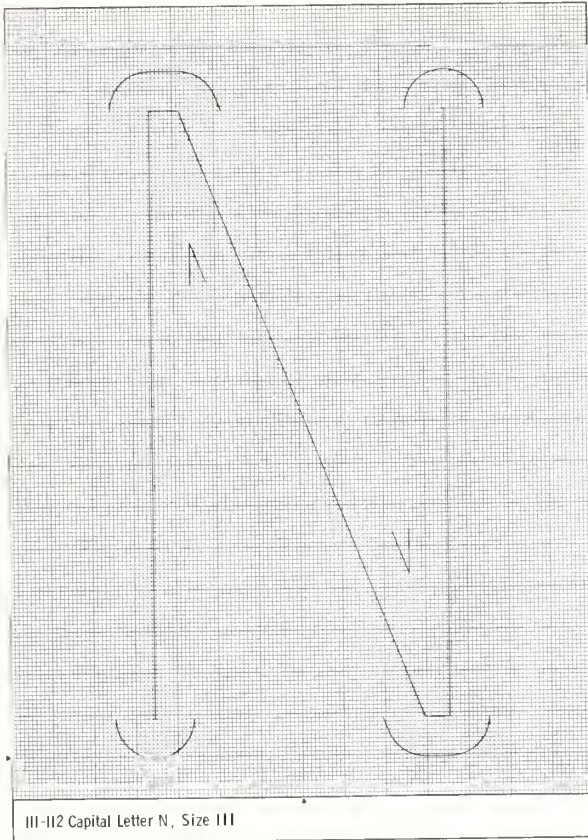




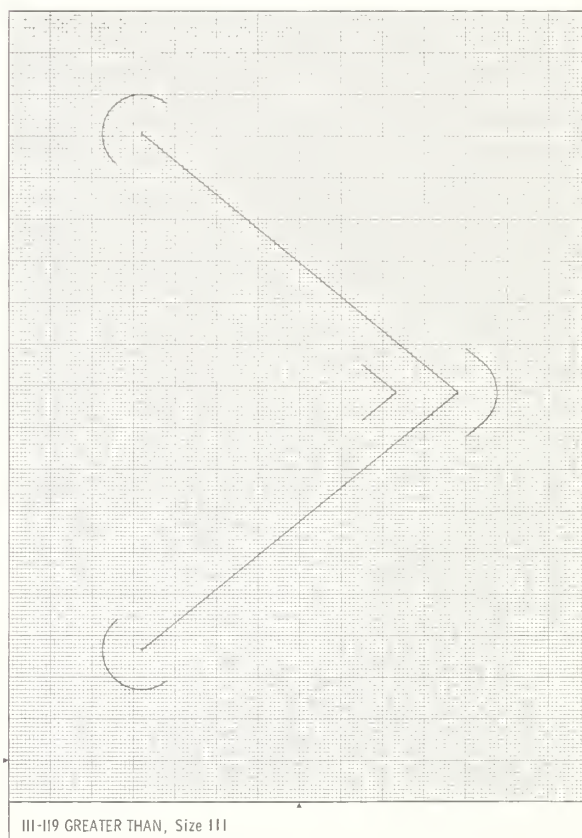
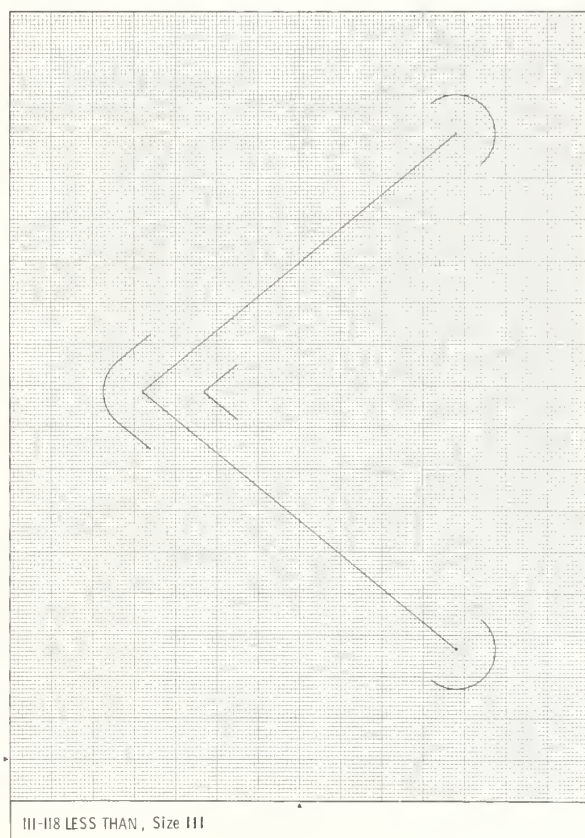
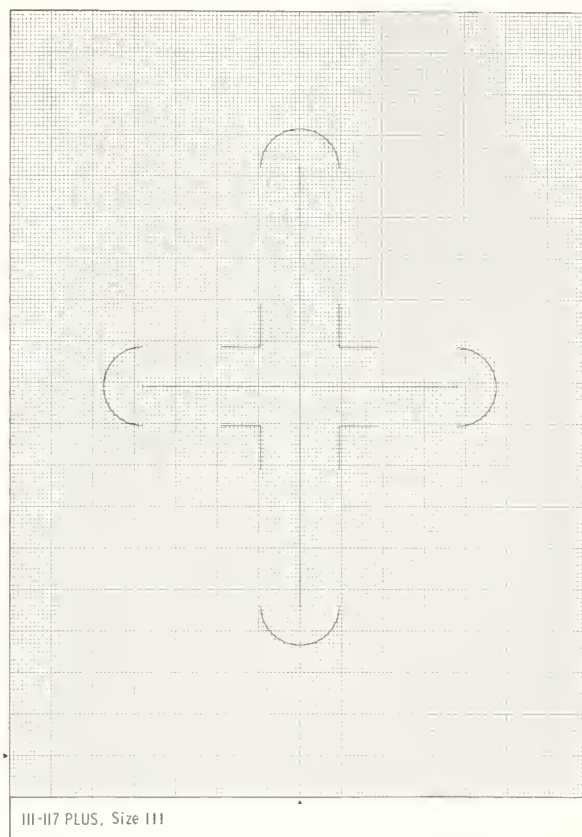
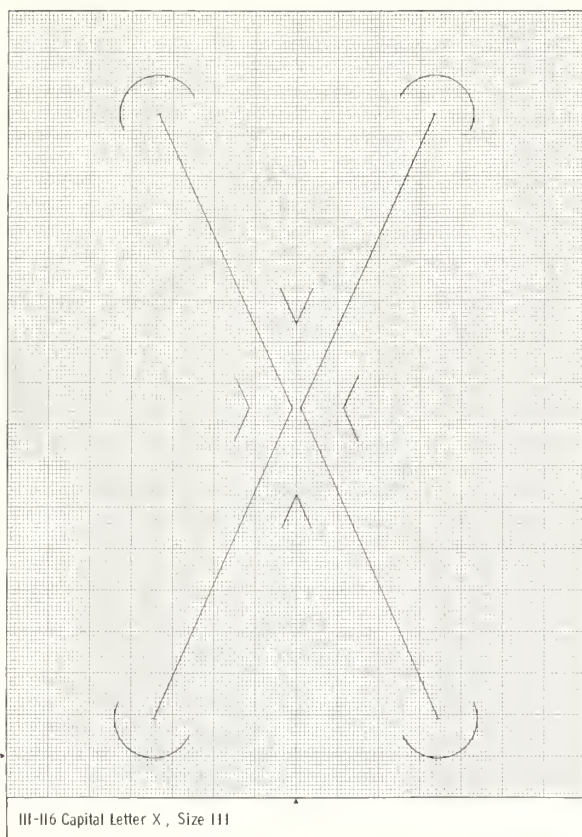




























OFFICIAL BUSINESS

PENALTY FOR PRIVATE USE, \$300

**If you're looking for a solution to  
your ADP standards problems . . .**



**DO YOU** need up-to-date information on Federal Standards in the computer/information processing field?

**MUST YOU** keep abreast of standards adopted under PL 89-306 (Brooks Bill)?

**THEN YOU** need to subscribe to NBS FIPS Pub series! (Federal Information Processing Standards Publications Series of the National Bureau of Standards)

**FIPS PUBS** contain Federal standards for hardware, software, applications and data. **FIPS PUBS** provide information on new and revised Federal standards as they become available.

**FIPS PUBS** are the official U.S. Government publications for Federal computer/information processing standards.

**FIPS PUBS ARE AVAILABLE ON A SUBSCRIPTION BASIS FROM THE U.S. GOVERNMENT PRINTING OFFICE.**

